Hegagi Mohamed Ali Ahmed

Necessary Conditions for Constrained Nonsmooth Fractional Optimal Control Problems

PhD thesis submitted to the Faculty of Science of University of Porto, Portugal, in fulfillment of the requirements for the degree of Doctor of Philosophy in Mathematics

Department of Mathematics
Faculty of Science, University of Porto, Portugal

2016

Necessary Conditions for Constrained Nonsmooth Fractional Optimal Control Problems

Hegagi Mohamed Ali Ahmed

Doutoramento em Matemática Aplicada

Departamento de Matemática, Universidade do Porto, Portugal.

2016

Orientador
Prof. Fernando Manuel Ferreira Lobo Pereira
Faculty of Engineering of the University of Porto, Portugal.

Co-orientador
Prof. Sílvio Marques de Almeida Gama
Faculty of Science of the University of Porto, Portugal.

PhD

Abstract

The main objective of this thesis is to provide a contribution to the body of results currently available in the optimal-control theory for dynamic control systems modeled by a fractional derivative of the state variable with respect to time. Therefore, the principal focus of this dissertation is on obtaining the smooth and nonsmooth necessary conditions of optimality in the form of a maximum principle of the Pontryagin type for fractional optimal-control problems.

The fractional optimal-control is a generalization of the corresponding integer optimal-control theories, which exhibit controlled system dynamics with arbitrary-order derivatives and integrals (usually, and without any loss of generality, of integer order).

The first aim of this thesis is to display a maximum principle of the Pontryagin type for the generalization of nonlinear fractional optimal-control problems, under proper assumptions with smooth data, and we propose an approach contributing to obtain a maximum principle concerning the recommended problem. The novelty of our approach consists in a more precise insight inherent to the use of optimal-control variational methods in the original modeling framework, which are quite distinct from those originated in the calculus of variation framework frequently used in the literature. Moreover, we apply the proposed necessary conditions to illustrate an example of generalization of fractional optimal-control problems, which is solved by the Mittag-Leffler function.

Before moving to the main goal of this thesis, *i.e.*, to derive a nonsmooth maximum principle for fractional optimal-control problems with state constraints, we extend fractional integrals to more general Radon measures. Being these results of independent interest, they play an important role in the derivation of the maximum principle with state constraints, since the associated adjoint multiplier encompass such a type of measures.

We also present a nonsmooth necessary condition of fractional differential inclusion with state constraints under weak assumptions, and we introduce important concepts and results concerning the existence and compactness of sets of the fractional trajectories.

Finally, we address the main goal of this dissertation, *i.e.*, to state a fractional optimal-control formulation, which requires very weak assumptions on the data of the problem, and exhibit important types of constraints, notably control constraints and state constraints, that have not been yet considered in the fractional context. The approach to the proof of the necessary conditions of optimality in the form of a maximum principle of Pontryagin type is based on penalization techniques, variational principles, nonsmooth analysis, and fractional calculus results. An example illustrating the application of the derived maximum principle is also included.

Resumo

O principal objetivo desta tese é proporcionar uma contribuição para o conjunto de resultados, atualmente existente, da teoria de controle ótimo de sistemas de controles dinâmicos modelados por derivadas fracionárias, nas variáveis de estado, em relação ao tempo. Neste sentido, o enfoque principal desta dissertação é estabelecer condições necessárias de otimalidade, suaves e não-suaves, sob a forma de um princípio de máximo do tipo de Pontryagin para problemas de controlo óptimo fracionários.

O controle ótimo fracionário é uma generalização da correspondente teoria de controle ótimo, em derivadas inteiras, onde a dinâmica do sistema é controlada por derivadas e integrais de ordem arbitrária (geralmente, e sem qualquer perda de generalidade, de ordem inteira).

O primeiro objetivo deste trabalho é apresentar um princípio do máximo, do tipo de Pontryagin, que genaraliza problemas de controle ótimo fracionários não-lineares, sob hipóteses adequadas, com dados suaves, e propomos uma abordagem que constrói um princípio do máximo. A novidade da nossa abordagem consiste numa visão mais precisa inerente ao uso de métodos variacionais de controlo ótimo no quadro da modelação original, que é muito distinto do que se obtém no campo do cálculo de variaçõese e, frequentemente, encontrados na literatura. Mais ainda, as condições necessárias que propomos, são ilustradas num exemplo que generaliza um conjunto de problemas de controle ótimo fracionários, que se resolvem graças ao uso da função de Mittag-Leffler.

Antes de passar ao objetivo principal desta tese, por outras palavras, estabelecer um princípio do máximo não-suave para problemas de controlo óptimo fracionário com restrições de estado, estendemos os integrais fracionários ao caso mais geral de medidas de Radon. Tendo este resultado o seu interesse próprio, ele desempenha um papel importante na derivação do princípio do máximo com restrições de estado, uma vez que o multiplicador adjunto associado engloba uma medida deste tipo.

Também apresentamos, sob hipóteses fracas, uma condição necessária não-suave de inclusão diferencial fracionária com restrições de estado e introduzimos conceitos importantes e resultados sobre a existência e compacidade dos conjuntos das trajetórias fracionárias.

Finalmente, abordamos o objetivo principal desta dissertação, isto é, estabelecer uma formulação de controle ótimo fracionário, a qual requer, nos dados do problema, hipóteses

muito fracas e exibem importantes tipos de restrições, nomeadamente, restrições de controlo e restrições de estado, as que ainda não foram consideradas no contexto fracionário. A demonstração das condições necessárias de otimalidade, sob a forma de um princípio do máximo do tipo Pontryagin, usa técnicas de penalização, princípios variacionais, análise não-suave e resultados de cálculo fracionário. Um exemplo, que ilustra a aplicação deste princípio do máximo que apresentamos, está também aqui incluído.

Acknowledgments

All appreciation and thanks to Allah who guided and helped me to achieve this thesis, and to all those who somehow contributed to the accomplishment of this PhD work.

First of all, I would like to express my most sincere gratitude to my supervisors, Prof. Fernando Manuel Ferreira Lobo Pereira, Full Professor in the Department of Electrical and Computer Engineering (DEEC, and SYSTEC) at the Faculty of Engineering (FEUP), University of Porto, Portugal, and Prof. Sílvio Marques de Almeida Gama, Associate Professor in the Department of Mathematics (and CMUP) at the Faculty of Science (FCUP) and Director of the Doctoral Course in Applied Mathematics, University of Porto, Portugal. I want to thank them primarily for accepting to jointly supervise this thesis, for their guidance, support, encouragement, accessibility, their remarks and comments during the meetings we held together, that helped me to understand the relevant concepts of the problem, and their friendship during my studies at the University of Porto. This thesis would not have been possible without their expertise.

I am grateful for the financial support given by Erasmus Mundus-Deusto University through the grant Erasmus Mundus Fatima Al Fihri Scholarship Program Lot 1 (EMA2 Lot 1) under grant agreement no. ALFI1201870. Also, I am thankful to the Portuguese coordinator, Ana Paiva, and to all the Erasmus Mundus team in Porto.

I wish to express my thankfulness to Dr. Marco Martins Afonso for his support and assistance during the last month of writing.

I thank some colleagues from the Mathematics Department of the University of Porto, especially Marcelo Trindade, Deividi Pansera, Muhammad Ali Khan, Teresa Daniela Grilo and Nikolaos Tsopanidis, for their support and assistance on mathematical questions and valuable insights.

Special thanks go to those friends that gave me the right motivation at the right time and make me feel that I am not alone. In particular, Mahmoud Mabrok, Hossameldeen Ahmed, Mohamed Alaa, Mohamed Al Atwany and his wife Nagwa Arafa.

I am grateful as well to all my friends here in Porto and also in Egypt, for their support, assistance, friendship over the past years, and for the entertainment that they provided to me. I would like to thank everybody who has been important to the successful realization of this thesis.

This humble effort is dedicated to my parents for their encouragement, support, and prayers for me, that always gave me strength in difficult times. I really thank you for the most important values you have given me, and I would like to thank my brothers, my sisters and all my family.

Last but no means least, I cannot thank enough my beloved wife and best friend, Elshimaa, for her love and encouragement that have stood beside me at all times, and who pushed me ahead with her unwavering faith. I would also like to acknowledge the most important persons in my life: my beloved son Mohamed and my daughter Menna.

Finally, I express my humble apologies to those whom I might have forgotten to thank.

Contents

Chapter 1

Introduction

1.1 Objectives

The goals of this thesis are the following:

a) Improvement of the fractional optimal-control (FOC) formulation and refinement of methods and approaches required for the developments of necessary conditions of optimality in the form of a Maximum Principle.

   This goal requires a new general formulation for fractional optimal-control problems (FOCPs), for which the objective function is given by an integral of fractional order, the velocity set is closed (it could even be a finite set of points), and a new approach to provide a necessary condition of optimality (a maximum principle of the Pontryagin type). The proposed approach is illustrated by an application example.

b) Formulation of optimal-control problems (OCPs) with fractional differential dynamics of incrementally increasing complexity, notably, in the presence of either control, endpoint, or state constraints, whose data features assumptions weaker than those usually considered for FOCPs, and more in line with the ones currently adopted for problems with integer derivatives.

   This goal requires not only the consideration of appropriate solution concepts, but also some investigation on the assumptions to be imposed in order to ensure the usefulness of the necessary conditions of optimality.

c) Refinement of methods and approaches — notably penalization techniques and variational methods — required to address the issues arising in item d) below, in the context of fractional differential control systems.

   This goal is an intermediate step required to address the remaining aims. It is clear that new approaches, with respect to the ones considered in the past for FOCPs, are required in order to deal with the challenges arising in the FOCP formulations to be considered, namely, new variational principles and new penalization techniques which, in turn, imply the need of nonsmooth calculus.

Fractional calculus (FC) will be exploited in order to investigate the new methods to be employed in the proofs of the results to be obtained.

d) Statement and proof of necessary conditions of optimality — a maximum principle of the Pontryagin type — for the case where we have fractional dynamics with state trajectories satisfying state constraints.

   To pursue this goal, several concepts of solutions for conventional OCPs are examined, and an essential role is played by the investigation of how the aspects intrinsic to the fractional nature of the derivatives in FOCPs affect these concepts. Moreover, it is expected that some aspects of the problem formulation, in particular, the assumptions on the data of the problem, might come into play when addressing these issues.

   The general approach to pursue this goal is to seek the application of Fermat's rule for a suitably defined infinite-dimensional optimization problem with constraints. This requires the use of specifically-defined penalization techniques coupled with a variational principle. Also, an example is included to illustrate the applicability of this approach.

1.2 Motivation

FOCPs are a generalization of classic OCPs for which either the dynamics of the control system is described by fractional differential equations, or the performance index is given by a fractional integration operator. The reason behind the use of fractional derivatives lies in the fact that they provide a more accurate description of the behavior of the considered dynamic system, and constitute an excellent tool for the characterization of memory and hereditary properties of several dynamic processes (see *e.g.*, Caponetto *et al.* [37], Das [48], Kilbas *et al.* [87], and Podlubny [125]).

OCPs have found applications in many different fields in real life, including biology (see *e.g.*, Behncke [29], Hawkins and Cornell [67], Jung *et al.* [79], and Lenhart and Workman [94]), ecology (see *e.g.*, Cohen [45], and Grigorieva *et al.* [64]), engineering (see *e.g.*, Torokhti and Howlett [145], and Zelikin and Borisov [156]), economics (see *e.g.*, Sethi and Thompson [136], and Zelikin and Borisov [156]), finance (see *e.g.*, Chen and Islam [40], and Davis and Elzinga [49]), resource allocation and management (see *e.g.*, Clark [41], and Sethi and Thompson [136]), medicine (see *e.g.*, Abello *et al.* [1], and Swan [137]), and so on.

In the last three decades, FC has attracted significant interest due to its multiple applications in many areas. In particular, fractional systems can be used in wide areas of application to indicate systems with long-range interactions or power-law memory (see *e.g.*, Laskin and Zaslavsky [92], and Luo and Afraimovich [99]). Fractional systems are often more appropriate than the usual ones (integer-order systems) in real-life applications, such as electrochemistry (Ichise *et al.* [72]), dielectric materials (Tarasov [138, 139]), viscoelastic materials (see *e.g.*, Bagley and Torvik [25, 26], and Renardy [130]), fractal networks (see *e.g.*, West *et al.* [152], and Arena *et al.* [10]), robotics (see *e.g.*, Gutiérrez *et al.* [65], Valério and Sá da Costa [148]), biological tissues (liver, valves, heart, brain, *etc.*; see, *e.g.*, Doehring *et al.* [52], Hoyt *et al.* [71], Kobayashi *et al.* [88], and Macé *et al.* [102]), electric circuits (see *e.g.*, Caponetto *et al.* [37], and Petras [124]), signal processing (see *e.g.*, Tseng and Lee [147], and Vinagre *et al.* [149]), control systems (see *e.g.*, Axtell and Bise [24], Caponetto *et al.* [37], Monje *et al.* [109], and Podlubny [126]), and so on.

In these years, some popular and simple structures have been proposed for fractional-order controllers, such as fractional-order Proportional-Integral-Derivative (PID) controllers, which are widely used in the industry (see *e.g.*, Hamamci [66], Podlubny [126], and Tavazoei [140]), and their subclasses, *i.e.* fractional-order PI and PD controllers (see *e.g.*, Jin *et al.* [74], and Luo and Chen [100]), as well as different generations of Commande Robuste Ordre Non Entier (CRONE) controllers which represent the first framework for non-integer-order system application in the automatic-control area (see *e.g.*, Lanusse *et al.* [90], Lanusse and Sabatier [91], and Oustaloup *et al.* [118], *etc.*). Moreover, fractional-order controllers have been employed in many useful applications, such as control of hard-disk drive servo systems (Luo *et al.* [101]), control of cement milling processes (Efe [54]), suppression of chaos in chaotic electrical circuits (Tavazoei *et al.* [141]), control of power electronic converters (Calderón *et al.* [36]), control of composite hydraulic cylinders (Zhao *et al.* [157]), control of irrigation canals (Feliu-Batlle *et al.* [57]), and many others.

The difference of complexity in defining and using calculus between the integer and fractional contexts together with the associated differences in the geometric interpretation, may explain why the development extents of the respective associated bodies of optimal-control theory differ so much. An overview on the research carried out on optimality conditions and methods for solving FOCPs reveals tremendous between both classes of problems. Similarly, huge gap exists for the extent of the development of existing optimal-control theory for conventional OCPs and FOCPs for which only performance integer integrals have been considered.

Starting with the pioneering work by Pontryagin and his team [127], in which a wide range of OCPs have been considered, notably, state constraints, discrete-time and the sophisticated concept of relaxed solution had already been exploited, conventional optimal-control theory went over the years through extremely-complicated developments, in which the assumptions of the problem were strongly weakened. Very diverse formulations (wide variety of constraints and also unbounded time horizons) and delicate issues of well posedness, sensitivity, nonsmoothness, and non-degeneracy have been considered, as it is easily attested in the works of many researchers, among which Pereira *et al.* [122], Arutyunov [11], Arutyunov *et al.* [14], Vinter [150], Clarke [42], and Clarke *et al.* [44].

Within the objectives of this thesis, steps towards contributions to close this gap are pursued. This effort requires the investigation of new methods that are centered on the development of variational results appropriated for FOCPs, as well as on the refinement of FC required to meet the demands of the challenges to be addressed.

1.3 Contributions

The first contribution of this thesis is the development of FOCPs by using a general formulation, where we employ the fractional integral operator in the cost function and describe the dynamics of the control system through the Caputo fractional derivative. Moreover, we introduce a new approach to prove necessary optimality conditions in the form of Pontryagin's maximum principle for the general formulation of fractional non-linear OCP under smooth assumptions on the data of the problem. Furthermore, we provide an example for this class of FOCP, we use the generalization of the Mittag-Leffler function to solve it, and compare our results with the classical ones (when $\alpha = 1$) to illustrate the effectiveness of these fundamental findings. The results of this chapter were presented as an abstract at the International Meeting AMS / EMS / SPM, 2015, Porto, Portugal. The complete version of the results has been published in [8].

The definition of fractional integral with respect to a general Radon (regular Borel) measure (designated by fractional Stieltjes integral) is also one of the contributions of this thesis, where this measure can be written as a decomposition of atomic (discrete), absolutely continuous, and singular continuous measures. We present some properties for this concept, too. These results will be relevant in the next two chapters, when we investigate the minimizer of the FOCP with state constraints. These results are submitted for publication.

Another contribution of this thesis is the construction of the necessary conditions for nonsmooth FOCPs, in which the dynamic system is characterized by a fractional differential inclusion with state constraints under some weak assumptions. These results will be relevant in the next chapter, in which we investigate the maximum principle for nonsmooth FOCPs with state constraints.

The final contribution deals with a new approach in the fractional context to prove the nonsmooth necessary conditions of optimality, in the form of a nonsmooth maximum principle for FOCP under weak assumptions on the data of the problem with state constraints. Here, we adopt the Jumarie fractional derivative to express the dynamics of the control system (they are defined for a continuous function not necessarily differentiable). Furthermore, we display an example for this type of fractional optimal-control problem with state constraint to show the effectiveness of our results. The results of this chapter are submitted for publication.

1.4 Organization

This thesis is organized as follows.

In Chapter 2, we talk about the state-of-the-art, present a brief overview of the fractional calculus, conventional optimal-control theory, and some methods to solve optimization problems with fractional differential equations.

In Chapter 3, we introduce an overview of the key results on conventional optimal-control theory, *i.e.*, the ones addressing problems whose dynamics is given by integer differential control systems. This includes the problem formulation, the necessary conditions of optimality — notably the ones in the form of a maximum principle — and OCP with constraints. Then, we pursue with developments on the optimal-control of fractional differential control systems. These include the problem formulation, necessary conditions of optimality, and numerical techniques to solve these problems.

In Chapter 4, we present a new approach to prove necessary conditions for optimality in the form of Pontryagin maximum principle for a general formulation of fractional non-linear OCPs, whose performance index is in the fractional integral form, and whose dynamics is given by a set of fractional differential equations (FDEs) in the Caputo sense. Moreover, we use a generalization of the Mittag-Leffler function to solve an example of this general formulation of FOCP, in order to illustrate the efficiency of our result. The results of this chapter have been accepted for publication in MMAS [8].

In Chapter 5, we present some important concepts of measure and integration theory, which help us understand the fractional Stieltjes integral, especially when we have a jump in the function. Furthermore, we find a formula for the fractional integral with respect to the measure in the Jumarie fractional integral sense. In this formula we are interested in the nonsmooth case, where the measure can be decomposed into an atomic and a non-atomic measure. Also, we present some properties of this formula.

In Chapter 6, we investigate the problem with dynamics given by a fractional differential inclusion, defined by a set valued map of the type $(t, x) \rightarrow F(t, x)$. We prove some important results related to differential inclusions (DI) in the fractional context, such as existence and compactness of fractional trajectories, and we formulate the nonsmooth necessary conditions for the fractional differential-inclusion problem with state constraints under certain weak assumptions.

In Chapter 7, we introduce a new procedure in the fractional context to provide necessary conditions in the form of a Pontryagin maximum principle for the FOCP, whose dynamics is given by the Jumarie fractional derivative subject to state constraints and under weak assumptions on the data of the problem. Moreover, we present an example to illustrate the effectiveness of our results.

Finally, in Chapter 8, we close this thesis by giving a summary of the present contributions from our works, and a few recommendations for future works.

Chapter 2

State-of-the-Art

In this chapter, we will introduce a brief review of key concepts and results of various domains – of fractional calculus, conventional optimal-control theory, and some methods to solve optimization problems with FDEs – which are relevant for the research developed in the thesis. Moreover, this overview will be also most helpful to appreciate the added value of this thesis with respect to the state of the art.

2.1 Fractional Calculus

Fractional calculus (FC) is a field of mathematics that deals with integrals and derivatives whose order may be an arbitrary real or complex number, thus generalizing the integer-order differentiation and integration. This field may be considered old and yet a quite young one. It is an old topic because its beginning can be traced back to Leibniz's letter to L'Hôpital in 1695, in which the notation for differentiation of non-integer order $\frac{1}{2}$ is discussed. Since then, fractional calculus has been developed gradually, being now one of the strongly researched areas of the mathematical analysis as attested by the number of publications.

The motivation for this lies in the increasing range of applications requiring the use of fractional differentiation and integration operators in various fields, notably, in pure and applied mathematics, physics, chemical, biological processes, engineering, economics, and control theory, among others (see *e.g.*, Hilfer *et al.* [70], Koh and Kelly [89], Mainardi [104], Makris *et al.* [105], and Rossikhin and Shitikova [133]). Thus, it can be considered a novel topic as well.

There are many definitions and several different approaches in fractional derivatives and integrals (see *e.g.*, De Oliveira and Tenreiro [51], Kilbas *et al.* [87], and Samko *et al.* [135]). Here, we introduce a review of some special functions that are used in FC as help tools, and some definitions for fractional integrals and derivatives of which we consider only the most useful ones for our purposes.

2.1.1 Special functions

Here, we present an overview of some definitions for special functions that were used along this thesis.

- Gamma function.

  One imperative function of the fractional calculus represents a continuous extension of the factorial function. That is, the Gamma function generalizes the factorial function to non-integer, negative and complex arguments. Also, called the Euler Gamma function, $\Gamma(z)$ is defined by

  $$\Gamma(z) = \int_0^\infty e^{-t} t^{z-1} dt,$$

  this integral being convergent for all complex $z \in \mathbb{C}, \Re(z) > 0$.

- Mittag-Leffler function and generalization.

  The Mittag-Leffler function plays a very important role in fractional calculus. It was first introduced in 1903 by the Swedish mathematician Gösta Mittag-Leffler (Mittag-Leffler [108]) and is given by

  $$E_\alpha(z) = \sum_{n=0}^{\infty} \frac{z^n}{\Gamma[n\alpha + 1]},$$

  where z is a complex variable, $\alpha > 0$ and $\Gamma(\cdot)$ is the Gamma function. It is called the one-parameter Mittag-Leffler function, as there is also a Mittag-Leffler function with two parameters in following form:

  $$E_{\alpha,\beta}(z) = \sum_{n=0}^{\infty} \frac{z^n}{\Gamma[n\alpha + \beta]} \quad \alpha, \beta > 0.$$

  In fact, the two-parameter Mittag-Leffler was introduced by Agarwal in 1953. The reason why now the two-parameter Mittag-Leffler is not called the Agarwal function, but simply the Mittag-Leffler function, is because Agarwal left the same notation as for the one-parameter Mittag-Leffler function (see *e.g.*, Das [48], Podlubny [125]). If $\beta = 1$, we have $E_{\alpha,1} = E_\alpha$, *i.e.*, the original one-parameter Mittag-Leffler function.

  The generalization of the Mittag-Leffler function is obtained by writing the argument in the form t^α. It is very important to solving FDEs, and defined as follows: Let $A \in \mathbb{R}^{n\times n}$, $\alpha > 0$, $\beta > 0$. Then, the generalization of the two-parameter Mittag-Leffler function is

  $$E_{\alpha,\beta}(At^\alpha) = \sum_{n=0}^{\infty} A^n \frac{t^{n\alpha}}{\Gamma[n\alpha + \beta]}.$$

If $\beta = 1$ we obtain the generalization of the one-parameter Mittag-Leffler function as

$$E_\alpha(At^\alpha) = \sum_{n=0}^{\infty} A^n \frac{t^{n\alpha}}{\Gamma[n\alpha + 1]}.$$

The generalization of the Mittag-Leffler function satisfies some interesting properties (see *e.g.*, Mozyrska and Torres [112], Prajapati [129]).

Proposition 2.1. *Let $\alpha > 0$ and $t \in [a, b]$. Then, the fractional derivative of the generalized Mittag-Leffler function obeys:*

$${}_a^C D_t^\alpha E_\alpha(A(t-a)^\alpha) \quad = \quad A E_\alpha(A(t-a)^\alpha),$$

where ${}_a^C D_t^\alpha$ is called the Caputo fractional operator, which is defined in the next section.

2.1.2 Fractional integrals and derivatives

There are many definitions and several different approaches for fractional derivatives and integrals (see Appendix A). Here, we introduce the fractional derivatives and integrals important for our work, such as the Caputo, Riemann-Liouville and Jumarie ones.

Definition 2.1. *Let $f(\cdot)$ be an integrable function in interval $[a, b]$. For $t \in [a, b]$ and $\alpha > 0$, the left and right Riemann-Liouville fractional integrals are, respectively, defined by*

$${}_a I_t^\alpha f(t) \quad := \quad \frac{1}{\Gamma(\alpha)} \int_a^t (t-\tau)^{\alpha-1} f(\tau) d\tau,$$

$${}_t I_b^\alpha f(t) \quad := \quad \frac{1}{\Gamma(\alpha)} \int_t^b (\tau-t)^{\alpha-1} f(\tau) d\tau,$$

where $\Gamma(\cdot)$ is the Euler Gamma function.

Definition 2.2. *Let $f(\cdot)$ be an absolutely continuous function in the interval $[a, b]$. For $\alpha > 0$, and $t \in [a, b]$, the left and right Riemann-Liouville fractional derivatives are, respectively, defined by*

$${}_a D_t^\alpha f(t) := \frac{d^n}{dt^n} \left({}_a I_t^{n-\alpha} f(t)\right) = \frac{1}{\Gamma(n-\alpha)} \left(\frac{d}{dt}\right)^n \int_a^t (t-\tau)^{n-\alpha-1} f(\tau) d\tau,$$

$${}_t D_b^\alpha f(t) := (-\frac{d}{dt})^n \left({}_t I_b^{n-\alpha} f(t)\right) = \frac{1}{\Gamma(n-\alpha)} \left(-\frac{d}{dt}\right)^n \int_t^b (\tau-t)^{n-\alpha-1} f(\tau) d\tau,$$

where $n \in \mathbb{N}$ is such that $n - 1 < \alpha \leq n$.

Definition 2.3. *Let $f(\cdot) \in AC^n[a,b]$. For $t \in [a,b]$ and $\alpha > 0$, the left and the right Caputo fractional derivatives are, respectively, defined by*

$$ {}_a^C D_t^\alpha f(t) := {}_a I_t^{n-\alpha} \frac{d^n}{dt^n} f(t) = \frac{1}{\Gamma(n-\alpha)} \int_a^t (t-\tau)^{n-\alpha-1} f^{(n)}(\tau) d\tau, $$

$$ {}_t^C D_b^\alpha f(t) := {}_t I_b^{n-\alpha} \left(-\frac{d}{dt}\right)^n f(t) = \frac{(-1)^n}{\Gamma(n-\alpha)} \int_t^b (\tau - t)^{n-\alpha-1} f^{(n)}(\tau) d\tau, $$

where $n \in \mathbb{N}$ is such that $n-1 < \alpha \leq n$.

If $\alpha = n \in \mathbb{N}_0$, then the Caputo and Riemann-Liouville fractional derivatives coincide with the ordinary derivative $\dfrac{d^n f(t)}{dt^n}$. For some properties and relations between the Caputo and Riemann-Liouville derivatives, see Appendix A.

FOCPs make use of different types of fractional derivatives. The most popular among them are the Caputo and Riemann-Liouville fractional derivatives, but both have some disadvantages. For example, the Caputo fractional derivative does not apply when the functions are not differentiable (*et al.* [132])), and the Riemann-Liouville fractional derivative of a constant is not equal zero.

To overcome this issue, there are a results proposed by Jumarie which involve a slightly modified definition of the fractional derivative of the Riemann-Liouville derivative (Jumarie [76–78]) in order to eliminate the disadvantages of the Riemann-Liouville and Caputo fractional derivatives. The Jumarie fractional derivative of a constant is equal to zero, and it is defined for a continuous (not necessarily differentiable) function.

Definition 2.4. *Let $f : [a,b] \to \mathbb{R}$ be a continuous function, $0 < \alpha < 1$ and $t \in [a,b]$. Then, the left and right Jumarie fractional derivatives $f_L^{(\alpha)}(t)$ and $f_R^{(\alpha)}(t)$ are, respectively, defined by*

$$ \begin{aligned} f_L^{(\alpha)}(t) &:= ({}_a D_t^\alpha [f(\cdot) - f(a)])(t) \\ &= \frac{1}{\Gamma(1-\alpha)} \frac{d}{dt} \int_a^t \frac{f(\tau) - f(a)}{(t-\tau)^\alpha} d\tau, \end{aligned} $$

$$ \begin{aligned} f_R^{(\alpha)}(t) &:= ({}_t D_b^\alpha [f(b) - f(\cdot)])(t) \\ &= -\frac{1}{\Gamma(1-\alpha)} \frac{d}{dt} \int_t^b \frac{f(b) - f(\tau)}{(\tau - t)^\alpha} d\tau. \end{aligned} $$

If $\alpha \geq 1$, we define

$$ f^{(\alpha)}(t) := \left(f^{(\alpha-n)}(t)\right)^{(n)}, \quad n \leq \alpha < n+1, \quad n \geq 1. $$

For the special case $0 < \alpha < 1$, *we have*

$$f^{(\alpha)}(t) = (f^{(\alpha-1)}(t))'.$$

Remark 2.1. *If* $f(a) = 0$, *then the left Jumarie fractional derivative coincides with the Riemann-Liouville fractional derivative.*

Definition 2.5. *Let* $f : [a, b] \to \mathbb{R}$ *be a continuous function and* $\alpha > 0$ *a real number in the interval* $[a, b]$. *The Jumarie fractional integral is defined by*

$$_aJ_t^\alpha f(t) = \frac{1}{\Gamma(\alpha+1)} \int_a^t f(\tau)(d\tau)^\alpha, \quad 0 < \alpha \leq 1.$$

Moreover, Jumarie introduced a fractional integral notation $(dt)^\alpha$, defined by

$$\int_a^t f(\tau)(d\tau)^\alpha = \alpha \int_a^t (t-\tau)^{\alpha-1} f(\tau) d\tau.$$

Important properties of the Jumarie fractional derivative and integral are shown in Appendix A.

Definition 2.6. *Let* $n - 1 < \alpha \leq n$. *The function* $f(\cdot)$ *is said to be an* $\alpha-$*absolutely continuous if satisfies*

$$f(t) = \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{\Gamma(k+1)} (t-a)^k + {_aI_t^\alpha} g(t), \quad t \in [a, b],$$

where $g(t) = {_aD_t^\alpha} f(t), \quad t \in [a, b]$.

For more details, see, *e.g.*, Hilfer [69], or Kilbas *et al.* [87].

2.1.3 Fractional integration by parts

Integration by parts plays an important role in deriving the generalized Euler-Lagrange equations for fractional variational problems, and in proving necessary optimality conditions for fractional optimal-control problems (see *e.g.*, Agrawal [4], Kilbas *et al.* [87], and Samko *et al.* [135]).

- Let $\varphi(t) \in L^p([a, b])$, $\psi(t) \in L^q([a, b])$, $1 \leq p \leq \infty$, $1 \leq q \leq \infty$ and $\frac{1}{p} + \frac{1}{q} \leq 1 + \alpha$ (we assume that $p > 1$ and $q > 1$, when $\frac{1}{p} + \frac{1}{q} = 1 + \alpha$), then,

$$\int_a^b \varphi(t)({_aI_t^\alpha} \psi)(t) dt = \int_a^b \psi(t)({_tI_b^\alpha} \varphi)(t) dt.$$

- Let $f(t) \in {}_aI_t^\alpha(L^p)$, $g(t) \in {}_tI_b^\alpha(L^q)$, where ${}_aI_t^\alpha(L^p)$, ${}_tI_b^\alpha(L^q)$ denote the ranges of the operators ${}_aI_t^\alpha$, ${}_tI_b^\alpha$ on L^p, L^q respectively. Then,

$$\int_a^b f(t)({}_aD_t^\alpha g)(t)dt = \int_a^b g(t)({}_tD_b^\alpha f)(t)dt,$$

and ${}_aI_t^\alpha(L^p)$, ${}_tI_b^\alpha(L^q)$ for any $1 \leq p \leq \infty$, $1 \leq q \leq \infty$, $\alpha > 0$ are defined by

$$\begin{aligned} {}_aI_t^\alpha(L^p) &:= \{f : f = {}_aI_t^\alpha \varphi, \qquad \varphi \in L^p(a,b)\}, \\ {}_tI_b^\alpha(L^q) &:= \{g : g = {}_tI_b^\alpha \psi, \qquad \psi \in L^q(a,b)\}. \end{aligned}$$

There is a formula for fractional integration by parts for the Caputo fractional derivatives, such that

$$\begin{aligned} \int_a^b f(t)({}_a^CD_t^\alpha g)(t)dt &= \int_a^b g(t)({}_tD_b^\alpha f)(t)dt + \sum_{i=0}^{n-1}({}_tD_b^{\alpha+i-n} f(t))g^{(n-1-i)}(t) \mid_{t=a}^{t=b}, \\ \int_a^b f(t)({}_t^CD_b^\alpha g)(t)dt &= \int_a^b g(t)({}_aD_t^\alpha f)(t)dt + \sum_{i=0}^{n-1}(-1)^{n+i}({}_aD_t^{\alpha+i-n} f(t))g^{(n-1-i)}(t) \mid_{t=a}^{t=b}, \end{aligned}$$

and in particular, for $\alpha \in (0,1)$, we have

$$\begin{aligned} \int_a^b f(t)({}_a^CD_t^\alpha g)(t)dt &= \int_a^b g(t)({}_tD_b^\alpha f)(t)dt + ({}_tI_b^{1-\alpha} f(t))g(t) \mid_{t=a}^{t=b}, \\ \int_a^b f(t)({}_t^CD_b^\alpha g)(t)dt &= \int_a^b g(t)({}_aD_t^\alpha f)(t)dt - ({}_aI_t^{1-\alpha} f(t))g(t) \mid_{t=a}^{t=b}. \end{aligned}$$

In the previous equations, when $\alpha \to 1$, we have the classical formula of integration by parts, (*i.e.*, $\int_a^b f(t)g'(t)dt = f(t)g(t) \mid_{t=a}^{t=b} - \int_a^b g(t)f'(t)dt$), because ${}_a^CD_t^\alpha = \frac{d}{dt}$, ${}_t^CD_b^\alpha = -\frac{d}{dt}$, ${}_aD_t^\alpha = \frac{d}{dt}$, ${}_tD_b^\alpha = -\frac{d}{dt}$, and ${}_aI_t^{1-\alpha}$, ${}_tI_b^{1-\alpha}$ are the identity operators.

Furthermore, Jumarie introduced a formula of integration by parts as follows

$$\begin{aligned} \int_a^b f^{(\alpha)}(t)g(t)(dt)^\alpha &= \int_a^b (f(t)g(t))^{(\alpha)}(dt)^\alpha - \int_a^b f(t)g^{(\alpha)}(t)(dt)^\alpha \\ &= \Gamma(\alpha+1)[f(t)g(t)]_a^b - \int_a^b f(t)g^{(\alpha)}(t)(dt)^\alpha. \end{aligned}$$

2.2 Overview on Conventional Optimal Control Theory

Classic OCPs arise naturally in many various fields and have been discussed for a long time, therefore a lot of work exists in the area of optimal-control of integer-order dynamic systems in engineering, science, economics, and many other fields. Thus, it is not surprising that a wide diversity of OCPs have been considered. The range of issues and problems include; (i) multiple types of the considered control variations which are related to the various types of minimum; (ii) types of constraints (state, control, mixed, isoperimetric, endpoint and intermediate state constraints, (iv) finite or infinite time-horizons, (v) sets of assumptions to avoid several types of non-degeneracies of the conditions; (vi) sensitivity results; (vii) robustness to perturbations and to unknown model parameters; (viii) different types of multipliers; and (ix) control measures (impulsive control), among others, (see *e.g.*, Arutyunov [11], Arutyunov *et al.* [13–18, 22], Arutyunov and Pereira [20, 21], Arutyunov *et al.* [12], Arutyunov *et al.* [19], Bryson [35], Clarke [42], Clarke *et al.* [44], Fraga and Pereira [59], Gamkrelidze [60], Dubovitskii and Milyutin [53], Gregory and Lin [63], Hestenes [68], Karamzin *et al.* [84, 85], Neustadt [115], Pereira and Silva [119–122], Pontryagin *et al.* [127], Vinter and Pereira [151], Pereira *et al.* [123], and Vinter [150]). In general, the specification of an OCP requires the following items: (i) state and control spaces; (ii) a performance index or a cost function, usually depending on the state and control variables, usually denoted by $x(\cdot)$ and $u(\cdot)$, respectively, to be minimized over the set of all admissible control processes; (iii) dynamic constraints that establish the relation between the state trajectory with the control input and an endpoint state variable value; and, possibly, (iv) one or more types of constraints such as pointwise control and/or state constraints, and joint control and state constraints. Above, by control process we mean any pair $(x; u)$ that satisfies the dynamic constraints. A control process is feasible or viable if it satisfies all the constraints, and it is optimal if provides a cost lower than that associated with any other feasible control process.

The main objective of OCPs is to determine the (open- or closed-loop) control strategy that optimizes (minimizes or maximizes) a given optimality criterion or performance index usually denoted by $J(\cdot)$. The performance index may be very general. It may simply be a function depending on the state and/or time (for free-time problems) endpoints, or also involve an integral whose integrand may be a function of the values of both the state and the control variables.

Remark that, since $\min\{J(\cdot)\} = -\max\{-J(\cdot)\}$, it is indifferent to consider either maximization (more often used in economics) or minimization (more often used in engineering). We will adopt the latter.

Let $J[x, u]$ be a performance index, $x \in X$ a set of state variables taking values in $\mathbb{R}^n$, a control variable u is a Borel measurable function, $\mathcal{U}$ a closed set of control variables such that $u \in \mathcal{U}$ taking values on some closed set $\Omega \subset \mathbb{R}^m$, $t \in [a, b]$, and $L(\cdot)$, $f(\cdot)$ and $g(\cdot)$ continuously differentiable functions in all three formulations. Remark that if $f(t, x, \Omega)$ is an open set for all (t, x), we have a calculus of variations problem. In what concerns the cost function, there are three main types of OCPs:

1. Bolza formulation

$$J[x, u] = g(a, x(a), b, x(b)) + \int_a^b L(t, x(t), u(t))dt,$$

2. Lagrange formulation

$$J[x,u] = \int_a^b L(t, x(t), u(t))dt,$$

3. Mayer formulation

$$J[x,u] = g(a, x(a), b, x(b)).$$

Subject to

$$\dot{x}(t) = f(t, x(t), u(t)), \quad x(a) = x_0.$$

Theorem 2.1. (Rodrigues *et al.* [131]). There are conditions under which these three optimal-control problems are equivalent.

2.2.1 Necessary optimality conditions

Consider the simplest nonlinear OCP that can be formulated as follows

$$\begin{aligned} \text{Minimize} \qquad & J[x,u] = \int_a^b L(t, x(t), u(t))dt, \\ \text{subject to} \qquad & \dot{x}(t) = f(t, x(t), u(t)), \\ & x(a) = x_0. \end{aligned}$$

In this OCP, $x(t) \in \mathbb{R}^n$ is the state variable, $u \in \mathcal{U}$ is the control variable, $L : [a,b] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is called the integrand, $f : [a,b] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ is a function defining the dynamics of the system for the case where the first-order derivative of the state variable x is considered, and the points a and b are called the initial and final time points, respectively. The pair $(x; u)$ is designated as control process. The hypotheses to be satisfied by $L(\cdot)$ and $f(\cdot)$ have to be such that a number of purposes can be considered in the investigation of the OCP. The simplest subset of purposes concerns the proper definition of the problem — existence and uniqueness of the solution of the differential equation for a given initial state and control function — and the necessary conditions of optimality are informative in the sense that they enable the successful reduction of the number of candidates to the solution of the problem.

When $\Omega = \mathbb{R}^m$ where Ω is the set of values taken by the control function, the necessary conditions of optimality can be written down by using Lagrange multiplier as follows:

$$\begin{aligned} \dot{x}(t) &= \frac{\partial H}{\partial \lambda}(t, x(t), u(t), \lambda(t)), \\ \dot{\lambda}(t) &= -\frac{\partial H}{\partial x}(t, x(t), u(t), \lambda(t)), \end{aligned}$$

$$0 = \frac{\partial H}{\partial u}(t, x(t), u(t), \lambda(t)),$$

where $\lambda(\cdot)$ is a Lagrange multiplier and $H(t, x, u, \lambda)$ is the Pontryagin function defined by

$$H(t, x, u, \lambda) = L(t, x, u) + \lambda^T f(t, x, u).$$

2.2.2 Pontryagin maximum principle

The maximum principle of optimal-control provides necessary conditions to be satisfied by the optimal-control process, the pair $(x; u)$. It has been an important tool in the many areas in which optimal-control plays a role. The well known Pontryagin Maximum Principle was developed in the mid 1950s in the Soviet Union by the Russian mathematician Lev Semenovich Pontryagin and his colleagues [127].

The Pontryagin maximum principle is stated as follows. Let $(x^*; u^*)$ be an optimal solution of the control problem

$$\begin{array}{ll} \text{Minimize} & J(x, u) = \int_a^b L(t, x(t), u(t))dt, \\ \text{subject to} & \dot{x}(t) = f(t, x(t), u(t)), \end{array}$$

where $x(a) = x_0$, $u \in \mathcal{U} \subset \mathbb{R}^m$, and $t \in [a, b]$. The Pontryagin function $H(t, x, u, \lambda)$ is defined as follows

$$H(t, x, u, \lambda) = L(t, x, u) + \lambda^T f(t, x, u),$$

where the functions $f : [a, b] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ and $L : [a, b] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ are, respectively, the state variable dynamics and the integrand of the cost functional. Both are differentiable with respect to (w.r.t.) x, Lebesgue-measurable w.r.t. t, and Borel-measurable w.r.t. u. Pontryagin's maximum principle states that the optimal input $u^*(\cdot)$ maximizes $H(t, x(t), u(t), \lambda(t))$ among all admissible inputs $u(\cdot)$, that is

$$H(t, x^*(t), u(t), \lambda^*(t)) \leq H(t, x^*(t), u^*(t), \lambda^*(t)),$$

for almost all (a.a.) $t \in [a, b]$, where $u^*(\cdot) \in \mathcal{U}$ is the optimal-control for the problem, $\lambda^*(\cdot)$ is the optimal co-state trajectory, and $x^*(\cdot)$ is the optimal state trajectory satisfying, respectively, the adjoint equation, $\dot{\lambda}^*(t) = -\frac{\partial}{\partial x} H(t, x^*(t), u^*(t), \lambda^*(t))$, and the state equations $\dot{x}^*(t) = \frac{\partial}{\partial \lambda} H(t, x^*(t), u^*(t), \lambda^*(t))$, with the boundary conditions $x^*(a) = x_0$ and $\lambda^*(b) = 0$.

When the final time is fixed and the Pontryagin function does not depend explicitly on time, then

$$H(x^*(t), u^*(t), \lambda^*(t)) = constant,$$

and, if the final time is free, then

$$H(x^*(t), u^*(t), \lambda^*(t)) = 0.$$

2.2.3 Optimal control problems with constraints

Constraints appear in different ways in OCPs. These constraints restrict the range of values of both the state and the control variables.

If the constraints are imposed on the control $u(\cdot)$ of the OCPs, they are called control constraints.

If the constraints are imposed only on the state trajectories of the OCPs $x(\cdot)$, they are called pure state constraints.

Finally, if the constraints are imposed on both the state and the control variables they are called mixed state constraints.

Let us show some types of the aforementioned constraints, imposed in the OCPs.

1. Control constraints

   The control $u \in \mathcal{U}$ is called control constraint, where $u(t)$ takes values in a closed set $\Omega(t)$ for almost every $t \in [a, b]$ and $\mathcal{U} : [a, b] \to \Omega(t) \subset \mathbb{R}^m$ is a multifunction taking on closed values.

2. Endpoint constraints

   The endpoint constraints can be imposed at the initial and /or terminal point(s) of a fixed time interval $[a, b]$, and the most general way to write them is

   $$(x(a), x(b)) \in C,$$

   where C is a closed set.

3. State constraints

   – Inclusion state constraints

     Let $X : [a, b] \to \mathbb{R}^n$ be a multifunction taking on closed values. Then, the inclusion state constraint is defined by

     $$x(t) \in X(t), \qquad \forall t \in [a, b].$$

   – Inequality state constraints

     Let $h : [a, b] \times \mathbb{R}^n \to \mathbb{R}$ be a given function. Then, an inequality state constraint is defined by

     $$h(t, x(t)) \leq 0, \qquad \forall t \in [a, b].$$

– Equality state constraints

Let $h : [a,b] \times \mathbb{R}^n \to \mathbb{R}$ be a given function. Then, an equality state constraint is defined by

$$h(t, x(t)) = 0, \qquad \forall t \in [a,b].$$

4. Mixed state constraints

– Inequality mixed state constraints

Let $h : [a,b] \times \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}^m$ be a given function. Then, an inequality mixed state constraint is defined by

$$h(t, x(t), u(t)) \leq 0, \qquad \text{a.e. } t \in [a,b].$$

– Equality mixed state constraints

Let $h : [a,b] \times \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}^m$ be a given function. Then, an equality mixed state constraint is defined by

$$h(t, x(t), u(t)) = 0, \qquad \text{a.e. } t \in [a,b].$$

2.2.4 Maximum principle

Here, we give an overview the nonsmooth maximum principle for OCPs with and without state constraints. In the 1970s Francis Clarke generalized the convex subdifferentials of Rockafellar to cover Lipschitz-continuous functions and, to some extent, lower semi-continuous functions (Clarke [43]). He applied nonsmooth analysis to optimization and optimal-control theory. Also in the 1970s, Mordukhovich proposed the idea of limiting subdifferentials, and demonstrated how transversality conditions in the nonsmooth maximum principle could be improved, thus, making the necessary conditions of optimality more precise.

We will consider the nonsmooth OCP with state constraints as the following

$$(P) \begin{cases} \text{Minimize} & g(x(a), x(b)) \\ \text{subject to} & \dot{x}(t) = f(t, x(t), u(t)), \quad \text{a.e. } t \in [a,b], \\ & h(t, x(t)) \leq 0, \quad \forall t \in [a,b], \\ & u \in \mathcal{U}, \quad \text{a.e. } t \in [a,b], \\ & (x(a), x(b)) \in C. \end{cases}$$

The problem (P) satisfies the following hypotheses:

(H1) the function $(t,u) \to f(t, x.u)$ is $\mathcal{L} \times \mathcal{B}-$measurable;

(H2) the function $f(t, ., u)$ is Lipschitz with a function K_f in L^1 for all $(t,u) \in \{(t, \Omega(t)) : t \in [a,b]\}$;

(H3) the graph of $\Omega(t)$ is $\mathcal{L}\times\mathcal{B}-$measurable, where $\mathrm{Gr}(\Omega)$ is the graph of the multifunction $\mathcal{U} : [a,b] \to \mathbb{R}^m$ defined by

$$\mathrm{Gr}(\Omega) := \{(t,u) \in [a,b] \times \mathbb{R}^m : u \in \Omega(t)\} ;$$

(H4) the function g is Lipschitz of rank K_g;

(H5) the function h is upper semicontinuous and for each $t \in [a,b]$ the function $h(t,\cdot)$ is Lipschitz with constant k_h.

The Pontryagin function $H : [a,b] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is defined by

$$H(t,x,p,u) = \langle p, f(t,x,u)\rangle .$$

Theorem 2.2. *(The Nonsmooth Maximum Principle for Optimal Control Problem (P) with State Constraints (Theorem 5.2.1; Clarke* [43]*).*

Let (x^*,u^*) be a strong local minimizer for the problem (P), and assume the previous hypotheses (H1)–(H5) are satisfied. Then, there exist an arc p, a scalar $\lambda \geq 0$, a non-negative Radon measure $\mu(\cdot)$ on $[a,b]$, and a measurable function $\gamma(\cdot)$ such that the following expressions are satisfied:

(i) The Non-triviality Condition

$$\|p\| + \|\mu\| + \lambda > 0,$$

(ii) the adjoint equation

$$-\dot{p}(t) \in \partial_x H(t, x^*(t), q(t), u^*(t), \lambda),$$

(iii) the maximum condition

$$H(t, x^*(t), q(t), u^*(t), \lambda) = max \{H(t, x^*(t), q(t), w, \lambda) : w \in \mathcal{U}(t)\} ,$$

(iv) the transversality condition

$$(p(a), -q(b)) \in \lambda\partial g(x^*(a), x^*(b)) + N_C(x^*(a), x^*(b)),$$

(v) $\gamma(t) \in \partial_x^{>} h(t, x^*(t))$, and μ is supported on the set

$$\{t : h(t, x^*(t)) = 0\}.$$

Here, $q(\cdot)$ (referred to in conditions (ii), (iii) and (iv)) is

$$q(t) = \begin{cases} p(t) + \int_{[a,t)} \gamma(s)\mu(ds), & t \in [a,b), \\ p(t) + \int_{[a,b]} \gamma(s)\mu(ds), & t = b, \end{cases}$$

and $\partial_x^{>}(\cdot)$ is a certain generalized gradient that takes as values certain subsets of the well known Clarke's generalized gradient, defined by

$$\partial_x^{>} h(t,x) := co\{\gamma = \lim_{i\to\infty} \gamma_i : \gamma_i \in \partial_x h(t_i, x_i), (t_i, x_i) \to (t,x), h(t_i, x_i) > 0 \ \forall i\}.$$

As it is clear from this result, there are a number of other objects inherent to the generalized nonsmooth calculus whose understanding requires an overview of several basic concepts. We provide an overview of these in what follows next.

Definition 2.7. (Clarke [43], Liu and Zeng [96]). Let a function $f\colon X \to R$ is Lipschitz near a given point x. The Clarke's generalized directional derivative of $f(\cdot)$ at the point $x \in X$ in the direction d defined by

$$f^{\circ}(x;d) := \lim_{\lambda\to 0} \sup_{y\to x} \frac{f(y+\lambda d) - f(y)}{\lambda},$$

where y is a vector in X and λ is a positive scalar.

The Clarke's subdifferential or generalized gradient of the function $f(\cdot)$ at the point $x \in X$, denoted by $\partial f(x)$ is a subset of X^*

$$\partial f(x) := \{x^* \in X^* \colon f^{\circ}(x;d) \geq \langle x^*, d\rangle, \forall d \in X\},$$

where X^* is a dual space of X.

Remark 2.2. *If the state constraints are absent, then either* $\gamma(\cdot) = 0$ *or the measure* $d\mu(\cdot) = 0$ *and the conditions become simpler.*

It is easy to see that under some circumstances, these conditions may degenerate. For example, if $C = \{x_0\} \times \mathbb{R}^n$, h is smooth with $h(t,x_0) = 0$, and $h(t,x(t)) < 0$ for all feasible $x(\cdot)$ with $t > t_0$, then is immediate to see that the multiplier $\lambda = 0$, $\gamma(0) = \nabla h(t_0, x_0)$ $d\mu = \delta_{t_0}(\cdot)$ and $p(0) = -\nabla h(t_0, x_0)$ satisfies the condition of the maximum principle of Pontryagin (in particular, is nontrivial) and, at the same time, does not give any information to select the extremals of the OCP.

This problem has been addressed by several authors with various approaches , being the more significant ones found in (see *e.g.*, Arutyunov *et al.* [14], and Dubovitskii1 and Milyutin [53]).

There are to key approaches to address this challenge: Either to impose additional conditions on the data of the problem like in (see *e.g.*, Vinter [150], and Arutyunov *et* *al.* [17,18]) , or resort to higher order information in order to make sure that the multiplier specified by the conditions of the maximum principle does not yield non-informative multipliers like in Mordukhovich [110].

2.3 Overview of some Methods to Solve Optimization Problems with Fractional Differential Equations

In this section, we review some recent papers in which some optimization problems with fractional differential constraints (OPFDC) (of which FOCPs are a special case) were solved.

While integer-order OCPs have been discussed for a long time, and a large body of theory and numerical techniques has been developed to solve them, the FOCPs constitute a new area with a limited number of publications and many open issues. A general formulation and a solution scheme for OPFDC were first introduced by Agrawal [3], where the OPFDC formulation was expressed using the fractional variational principle and the Lagrange-multiplier technique, and the fractional dynamics of the OPFDC is defined in terms of the Riemann-Liouville fractional derivatives. He considered the state and the control variables as linear combinations of test functions, and dealt with the linear quadratic optimal-control problem as follows. Let the functions $q(t) \geq 0$ and $r(t) > 0$, then

$$\begin{aligned} (P)\,\text{Minimize} \qquad & J(u) = \frac{1}{2}\int_0^1 [q(t)x^2(t) + r(t)u^2(t)]dt, && (2.1)\\ \text{subject to} \qquad & {}_0D_t^{\alpha}x(t) = a(t)x(t) + b(t)u(t), && (2.2) \end{aligned}$$

where $x(0) = x_0$ and $\alpha \in (0,1)$. The necessary conditions of optimality for (P), in the Euler-Lagrange form, lead to a system of equations composed by (2.2) and

$$\begin{aligned} 0 &= r(t)u(t) + b(t)\lambda(t), && (2.3)\\ {}_tD_1^{\alpha}\lambda(t) &= q(t)x(t) + a(t)\lambda(t). && (2.4) \end{aligned}$$

The control variable $u(t)$ is obtained by using (2.3) as a function of the unknown co-state variable $\lambda(\cdot)$. In order to determine both the state and co-state functions, this equation together with equation (2.2) yields

$${}_0D_t^{\alpha}x(t) = a(t)x(t) - r^{-1}(t)b^2(t)\lambda(t). \tag{2.5}$$

Thus, the control function is determined by jointly solving (2.4) and (2.5) subject to the terminal conditions $x(0) = x_0$, $\lambda(1) = 0$. Agrawal used an approximate numerical method to find $x(t)$ and $\lambda(t)$, by using the shifted Legendre polynomials

$$P_j(t) = (-1)^j \sum_{k=0}^{j} \binom{j}{k} \binom{j+k}{k} (-t)^k,$$

that satisfy the following orthonormality conditions:

$$\int_0^1 P_j(t) P_k(t) dt = \delta_{jk} = \begin{cases} 0, & j = k, \\ 1, & j \neq k, \end{cases}$$

(here δ_{jk} is the Kronecker delta function). After some calculations and simplifications, we obtain the following system of $2m+2$ equations in $2m+2$ unknowns:

$$\begin{aligned}
0 &= \sum_{k=1}^{m} [F_1(j,k) - F_2(j,k)] c_k + \sum_{k=1}^{m} F_3(j,k) d_k + P_j(0)\mu_1, \\
0 &= -\sum_{k=1}^{m} [F_{0x}(j,k) c_k + \sum_{k=1}^{m} [F_4(j,k) - F_2(j,k)] d_k + P_j(1)\mu_2, \\
x_0 &= \sum_{k=1}^{m} P_k(0) c_k, \\
0 &= \sum_{k=1}^{m} P_k(1) d_k,
\end{aligned}$$

where μ_1 and μ_2 are the Lagrange multipliers associated with the terminal conditions, and $F_{0x}(j,k)$ and $F_1(j,k)$ to $F_4(j,k)$ are defined as:

$$\begin{aligned}
F_{0x}(j,k) &= \int_0^1 q(t) P_j(t) P_k(t) dt, \\
F_1(j,k) &= \int_0^1 P_j(t) {}_0D_t^{\alpha} P_k(t) dt, \\
F_2(j,k) &= \int_0^1 a(t) P_j(t) P_k(t) dt, \\
F_3(j,k) &= \int_0^1 r^{-1}(t) b^2(t) P_j(t) P_k(t) dt, \\
F_4(j,k) &= \int_0^1 P_j(t) {}_tD_1^{\alpha} P_k(t) dt.
\end{aligned}$$

An approximate solution to this problem is obtained by linear combinations of the shifted Legendre polynomials.

A direct numerical technique to solve OPFDC was used by Agrawal and Baleanu [5], where the authors consider a Hamiltonian formulation. They consider the following OPFDC, find the optimal-control $u(\cdot)$ that minimizes the performance index

$$J(u) = \int_0^1 f(x(t), u(t), t)dt,$$

subject to the system dynamic constraints

$${}_0D_t^\alpha x(t) = g(x(t), u(t), t),$$

with the initial condition $x(0) = x_0$, where $x(t)$ is the state variable, $f(\cdot)$ and $g(\cdot)$ are two given functions, and $0 < \alpha < 1$. Let $\lambda(\cdot)$ be the Lagrange multiplier. The Hamiltonian of the system is given by

$$H(x(t), u(t), \lambda(t), t) = f(x(t), u(t), t) + \lambda(t)g(x(t), u(t), t).$$

Then, the necessary conditions in terms of a Hamiltonian for the OPFDC are given by

$$\begin{aligned} {}_tD_1^\alpha \lambda(t) &= \frac{\partial H}{\partial x}, \\ 0 &= \frac{\partial H}{\partial u}, \\ {}_0D_t^\alpha x(t) &= \frac{\partial H}{\partial \lambda}, \end{aligned}$$

subject to the endpoint conditions $x(0) = x_0$, and $\lambda(1) = 0$. The authors focus on the problems with quadratic performance index like in equations (2.1) to (2.5), and they use a direct numerical method to compute $x(\cdot)$ and $\lambda(\cdot)$, by adopting the Grünwald-Letnikov definition. In this method, the entire time domain is organized into N equal domains, labeled by $0, 1, \ldots, N$. Then, the time at node j is given by $t_j = jh$, where $h = \frac{1}{N}$. By using the Grünwald-Letnikov concept, equation (2.5) at node i can be approximated as (see *e.g.*, Kilbas *et al.* [87], and Podlubny [125])

$$\frac{1}{h^\alpha} \sum_{j=0}^{i} w_j^{(\alpha)} x_{i-j} = a(ih)x_i - r^{-1}(ih)b^2(ih)\lambda_i, \qquad i = 1, \ldots, N,$$

where x_i and λ_i are the numerical approximations of $x(\cdot)$ and $\lambda(\cdot)$ at node i, and $w_j^{(\alpha)}$, $j = 0, \ldots, i$, are the coefficients defined by

$$w_j^{(\alpha)} = (-1)^j \binom{\alpha}{j}.$$

Similarly, equation (2.4) at node i can be approximated as

$$\frac{1}{h^\alpha}\sum_{j=i}^{N} w_{j-i}^{(\alpha)}\lambda_j = q(ih)x_i + a(ih)\lambda_i, \qquad i = 0, \ldots, N-1.$$

These equations provide a system of $2N$ equations in $2N$ unknowns, that can be solved by using various schemes such as a direct Gaussian elimination. From this system we can get $x(t)$ and $\lambda(t)$, and substitute in (2.3) to obtain $u(t)$.

Agrawal *et al.* [6] used the above-mentioned idea — Hamiltonian formulations for OPFDC and fractional derivative Grünwald-Letnikov concept — to solve the FDE involving the state and control variables. However, they considered OPFDC with vector-valued state and control variables. They state the OPFDC as follows

$$\begin{aligned} &\text{Minimize} \qquad J(u) = \frac{1}{2}\int_0^1 [q(t)x^2(t) + r(t)u^2(t)]dt, \\ &\text{subject to} \qquad {}_0D_t^\alpha x(t) = a(t)x(t) + b(t)u(t), \end{aligned}$$

with endpoint conditions $x(a) = c$ and $x(b) = d$. Here, the state and control variables, $x(t)$ and $u(t)$, are respectively n_x and n_u vectors, f and g are respectively a scalar and a n_x vector functions, and c and d are given vectors. The dimensions n_x and n_u satisfy the relation $n_u \leq n_x$.

Another slightly different variation of the previous method was introduced by Baleanu *et al.* [27]. The authors considered the same problem but used modified Grünwald-Letnikov approximations for left and right fractional derivatives

$$\begin{aligned} {}_0D_t^\alpha x(t_{i-\frac{1}{2}}) &\cong \frac{1}{h^\alpha}\sum_{j=0}^{i} w_j^{(\alpha)} x_{i-j}, \quad i = 1, \ldots, n, \\ {}_tD_1^\alpha x(t_{i+\frac{1}{2}}) &\cong \frac{1}{h^\alpha}\sum_{j=0}^{n-i} w_j^{(\alpha)} x_{i+j}, \quad i = n-1, n-2, \ldots, 0, \end{aligned}$$

where $w_j^{(\alpha)}$ are the coefficients satisfying $w_0^{(\alpha)} = 1$ and $w_j^{(\alpha)} = \left(1 - \frac{\alpha+1}{j}\right) w_{j-1}^{(\alpha)}, j = 1, \ldots, n$. These approximations are carried out at the central points of a certain discretization of the time horizon. Then, the time domain $[0, 1]$ is divided into n equal parts, and the fractional derivatives ${}_0D_t^\alpha x$ and ${}_tD_1^\alpha \lambda$ are approximated at the center of each segment, $x(t_{i-\frac{1}{2}})$ being defined as the average of the two end values of the segment:

$$x\left(t_{i-\frac{1}{2}}\right) = \frac{x_{i-1} + x_i}{2}.$$

Similar approximations are considered for $x(t_{i+\frac{1}{2}})$, $\lambda(t_{i-\frac{1}{2}})$ and $\lambda(t_{i+\frac{1}{2}})$.

By substituting these approximations in (2.4) and (2.5), the following system of equations is obtained:

$$\begin{aligned}
\frac{1}{h^\alpha}\sum_{j=0}^{i} w_j^{(\alpha)} x_{i-j} &= \frac{1}{2}a(i_1h)(x_{i-1}+x_i) - \frac{1}{2}r^{-1}(i_1h)b^2(i_1h)(\lambda_{i-1}+\lambda_i), \quad i=1,\dots,n, \\
\frac{1}{h^\alpha}\sum_{j=i}^{n-i} w_j^{(\alpha)} \lambda_{i+j} &= \frac{1}{2}q(i_2h)(x_{i+1}+x_i) + \frac{1}{2}a(i_2h)(\lambda_{i-1}+\lambda_i), \quad i=n-1,\dots,0,
\end{aligned}$$

where $i_1 = i-\frac{1}{2}$ and $i_2 = i+\frac{1}{2}$. This system of $2n$ linear equations in terms of $2n$ unknowns can be solved using a standard linear solver.

A numerical technique based on the Legendre orthonormal polynomial is the basis for solving OPFDC as discussed by Lotfi *et al.* [98]. The authors focus on OCPs with the quadratic performance index, and the fractional dynamics is defined in terms of the Caputo fractional derivatives, the solution method being based on the Legendre orthonormal polynomial approximation without using Hamiltonian conditions. They consider the following OPFDC formulation:

$$\begin{aligned}
(P)\ \text{Minimize} \quad & J(u) = \frac{1}{2}\int_0^1 [q(t)x^2(t) + r(t)u^2(t)]dt, \\
\text{subject to} \quad & {}_0D_t^\alpha x(t) = a(t)x(t) + b(t)u(t), \\
& x(0) = x_0,
\end{aligned}$$

where $q(t) \geqslant 0$, $r(t) > 0$ and $b(t) \neq 0$.

The state and control variables, $x(\cdot)$ and $u(\cdot)$, are expanded by using the Legendre basis $\Psi(\cdot)$ as follows:

$$\begin{aligned}
x(t) &\simeq (C^T I^\alpha + d^T)\Psi(t), \\
u(t) &\simeq U^T \Psi(t),
\end{aligned}$$

where I^α is the matrix operator of fractional integration of order α, $C^T = [c_0,\dots,c_m]$, $U^T = [u_0,\dots,u_m]$ and $d^T = [x_0, 0, \dots, 0]$.

After some calculations, the performance index J can be approximated by

$$\begin{aligned}
J \simeq\ & J[C,U] \\
=& \frac{1}{2}\int_0^1 [(Q^T\Psi(t))((C^T I^\alpha + d^T)\Psi(t)\Psi(t)^T(C^T I^\alpha + d^T)^T) + (R^T\Psi(t))(U^T\Psi(t)\Psi(t)^T U)]dt,
\end{aligned}$$

and the dynamical system as

$$C^T\Psi(t) - A^T\Psi(t)\Psi(t)^T(C^T I^\alpha + d^T)^T - B^T\Psi(t)\Psi(t)^T U = 0.$$

Then, the authors show that the dynamical-system approximation can be converted into a linear system of algebraic equations of the form

$$C^T - (C^T I^\alpha + d^T)^T \tilde{V} - U^T \tilde{W} = 0,$$

for some operators $\tilde{V}$ and $\tilde{W}$.

Now, define

$$J^*[C, U, \lambda] = J[C, U] + [C^T - (C^T I^\alpha + d^T)^T \tilde{V} - U^T \tilde{W}]\lambda,$$

where $\lambda^T = [\lambda_0, \lambda_1, \ldots, \lambda_m]$, $\tilde{V} = [\tilde{v}_{ij}]_{1 \leq i, j \leq m+1}$ and $\tilde{W} = [\tilde{w}_{ij}]_{1 \leq i, j \leq m+1}$. Then, the application of the necessary conditions of optimality yields

$$\begin{aligned}
\frac{\partial J^*}{\partial C} &= 0, \\
\frac{\partial J^*}{\partial U} &= 0, \\
\frac{\partial J^*}{\partial \lambda} &= 0,
\end{aligned}$$

where by $\frac{\partial J^*}{\partial C} = 0$ the system $\frac{\partial J^*}{\partial C_j} = 0, \quad j = 0, \ldots, m$ is meant. These equations can be solved for C, U and λ by using the Newton iterative method.

After computing C and U, the approximate values of $u(t)$ and $x(t)$ are obtained with the equations of the Legendre expansion. Some properties of Legendre polynomials and the convergence are also considered in this work (see Lotfi *et al.* [98]).

Another slightly different approach is discussed by Yousefi *et al.* [155]. In this work, the Hamiltonian conditions, the Riemann-Liouville fractional derivative and the Legendre multiwavelet collocation method are used to solve the OPFDC. The OPFDC is stated as follows:

$$\begin{aligned}
\text{Minimize} \qquad & J(u) = \int_{t_0}^{t_f} f(x(t), u(t), t)dt, \\
\text{subject to} \qquad & {}_0D_t^\alpha x(t) = g(x(t), u(t), t),
\end{aligned}$$

where $x(t_0) = x_0$, $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^m$, f and g are respectively a scalar and a n vector-valued functions. Then, the necessary conditions in terms of a Pontryagin function for the OPFDC are given by:

$${}_tD_{t_f}^\alpha \lambda(t) = \frac{\partial H}{\partial x},$$

$$\begin{aligned} 0 &= \frac{\partial H}{\partial u}, \\ {}_{t_0}D_t^{\alpha} x(t) &= \frac{\partial H}{\partial \lambda}, \\ x(t_0) &= x_0, \qquad \lambda(t_f) = 0. \end{aligned}$$

The adopted solution method is based on the approximation of $x(\cdot)$, $u(\cdot)$ and $\lambda(\cdot)$ by a truncated series of Legendre multiwavelets for $t \in [t_0, t_f]$ as follows:

$$\begin{aligned} x(t) &\simeq \sum_{i=0}^{2^k-1} \sum_{j=0}^{M} (t - t_0) c x_{ij} \Psi_{ij}(t) + x_0, \\ u(t) &\simeq \sum_{i=0}^{2^k-1} \sum_{j=0}^{M} c u_{ij} \Psi_{ij}(t), \\ \lambda(t) &\simeq \sum_{i=0}^{2^k-1} \sum_{j=0}^{M} (t - t_f) c \lambda_{ij} \Psi_{ij}(t), \end{aligned}$$

where, for $\frac{n(t_f - t_0)}{2^k} + t_0 \leq t \leq \frac{(n+1)(t_f - t_0)}{2^k} + t_0$,

$$\Psi_{nm}(t) = \sqrt{2m+1} \frac{2^{\frac{k}{2}}}{\sqrt{t_f - t_0}} P_m \left(\frac{2^k (t - t_0)}{t_f - t_0} - n \right),$$

$P_m(\cdot)$ being shifted Legendre polynomials. After using the collocation points P_i, such that $1 \leq i \leq 2^k(M+1)$, as the roots of Chebyshev polynomials of degree $2^k(M+1)$, the following algebraic system of equations is obtained:

$$\begin{aligned} F_1(x(P_i), u(P_i), \lambda(P_i)) &= 0, \\ F_2(x(P_i), u(P_i), \lambda(P_i)) &= 0, \\ F_3(x(P_i), u(P_i), \lambda(P_i)) &= 0, \end{aligned}$$

where $F_1(x(t), u(t), \lambda(t))$, $F_2(x(t), u(t), \lambda(t))$ and $F_3(x(t), u(t), \lambda(t))$ correspond to the equations of the necessary conditions of optimality. After solving these equations, the coefficients of the series by Legendre multiwavelets that approximate $x(t)$, $u(t)$ and $\lambda(t)$ are obtained. The well-known Chebyshev polynomials on the interval $[t_0, t_f]$ can be determined by the following recurrence formula:

$$T_{n+1}(t) = 2\left(\frac{2t}{t_f - t_0} - \frac{t_0 + t_f}{t_f - t_0}\right) T_n(t) - T_{n-1}(t),$$

with $T_0(t) = 1$ and $T_1(t) = \frac{2t}{t_f - t_0} - \frac{t_0 + t_f}{t_f - t_0}$.

Another idea for solving OPFDC, introduced by Jelicic and Petrovacki [73], consists in transforming the fractional problem into a classic integer-order problem by using an expansion formula for fractional derivatives. The work is based on an approximation formula obtained by Atanackovic and Stankovic [23]. The authors state the OPFDC as follows:

$$\text{Minimize} \qquad J(u) = \int_0^1 f(x(t), u(t), t)dt,$$

$$\text{subject to} \qquad \dot{x}(t) + k\, {}_0D_t^\alpha x(t) = g(x(t), u(t), t),$$

where $x(t_0) = x_0$ and k is a given constant. The necessary conditions of optimality for this problem are

$$\begin{aligned} \dot{x}(t) &= -k\, {}_0D_t^\alpha x(t) + g(x(t), u(t), t), \\ 0 &= \frac{\partial f}{\partial u} + \lambda(t)\frac{\partial g}{\partial u}, \\ \dot{\lambda}(t) &= k\, {}_tD_1^\alpha \lambda(t) - \lambda(t)\frac{\partial g}{\partial x}, \end{aligned}$$

subject to $x(0) = x_0$ and $\lambda(1) = 0$. After some calculations, the approximation

$${}_0D_t^\alpha x(t) \approx A(\alpha)t^{-\alpha}x(t) + \sum_{p=2}^{N} B(\alpha, p)t^{1-\alpha-p}V_p(t),$$

is used, where

$$\begin{aligned} A(\alpha) &= \frac{1}{\Gamma(1-\alpha)} - \frac{1}{\Gamma(1-\alpha)\Gamma(2-\alpha)} \times \sum_{p=2}^{N} \frac{\Gamma(p-1+\alpha)}{(p-1)!}, \\ B(\alpha, p) &= -\frac{1}{\Gamma(1-\alpha)\Gamma(2-\alpha)} \times \frac{\Gamma(p-1+\alpha)}{(p-1)!}, \\ V_n(x^{(p)})(t) &= \int_0^t x^{(p)}(\tau)\tau^n d\tau, \qquad n \in \mathbb{N}, \qquad t \geq 0, \end{aligned}$$

denoting by $V_n(x^{(p)})(t)$, $n \in \mathbb{N}$, the n^{th} moment of the function $x^{(p)}(\cdot)$, and by $x^{(p)}(\cdot)$, $p \in \mathbb{N}$, the p^{th} derivative of x. Then, the problem becomes a classic integer-order problem, and the fractional system is transformed into the following ordinary system:

$$\text{Minimize} \qquad J(u) = \int_0^1 f(t, x(t), u(t))dt,$$

$$\text{subject to} \qquad \begin{cases} \dot{x}(t) = -k(A(\alpha)t^{-\alpha}x(t) - \sum_{p=2}^{N} B(\alpha, p)t^{1-\alpha-p}V_p(t)) + \quad f(t, x(t), u(t)), \\ \\ \dot{V}_p(t) = (1-p)(t-0)^{p-2}x(t), \end{cases}$$

with the endpoint conditions $V_p(0) = 0$ for $p = 2, \ldots, N$, and $x(0) = x_0$. This problem can be solved by means of classical optimal-control theory, since fractional derivatives do not appear in its formulation.

Another similar procedure is discussed by Pooseh *et al.* [128]. The authors adopt the Caputo fractional derivative to model the dynamic constraints, and use an approximation formula to convert the fractional problem into an integer-order one with free terminal time. The OPFDC is stated as follows:

$$\begin{aligned} \text{Minimize} \qquad & J(u) = \int_a^T L(t, x(t), u(t))dt + \phi(T, x(T)), \\ \text{subject to} \qquad & M\dot{x}(t) + N\, {}_a^C D_t^\alpha x(t) = f(t, x(t), u(t)), \end{aligned}$$

with the endpoint condition $x(a) = x_a$. Here, M and N are non-zero and x_a is a fixed real number.

The optimal triplet (x, u, T) must satisfy the necessary conditions of optimality which are given in the next theorem.

Theorem 2.3. *If (x, u, T) minimizes the performance index while satisfying the dynamic constraints with the boundary condition, then there exists a function $\lambda(\cdot)$ for which the following conditions hold:*

- *The Hamiltonian system*

$$\begin{cases} M\dot{x}(t) + N\, {}_a^C D_t^\alpha x(t) = \frac{\partial H}{\partial \lambda}(t, x(t), u(t), \lambda(t)), & t \in [a, T], \\ \\ M\dot{\lambda}(t) - N\, {}_t D_T^\alpha \lambda(t) = -\frac{\partial H}{\partial x}(t, x(t), u(t), \lambda(t)), & t \in [a, T]. \end{cases}$$

- *The stationary condition*

$$\frac{\partial H}{\partial u}(t, x(t), u(t), \lambda(t)) = 0.$$

- *The transversality conditions*

$$[H(t, x(t), u(t), \lambda(t)) - N\lambda(t) {}_a^C D_t^\alpha x(t) + N\dot{x}(t)\, {}_t I_T^{1-\alpha} \lambda(t) + \frac{\partial \phi}{\partial t}(t, x(t))]_{t=T} = 0,$$

$$[M\lambda(t) + N {}_t I_T^{1-\alpha} \lambda(t) + \frac{\partial \phi}{\partial x}(t, x(t))]_{t=T} = 0,$$

where $H(t, x(t), u(t), \lambda(t))$ is the Pontryagin function given by

$$H(t, x(t), u(t), \lambda(t)) = L(t, x(t), u(t)) + \lambda(t) f(t, x(t), u(t)).$$

The authors use the expansion formula for fractional derivatives and the relation between Riemann-Liouville and Caputo derivatives to transform the given OPFDC into a classic integer-order optimal optimization problem as follows. The left Riemann-Liouville fractional derivative is approximated by

$$ {}_aD_t^\alpha x(t) \approx A(\alpha,N)(t-a)^{-\alpha}x(t) + B(\alpha,N)(t-a)^{1-\alpha}\dot{x}(t) - \sum_{p=2}^{N} C(\alpha,p)(t-a)^{1-\alpha-p}V_p(t), $$

where $V_p(t)$ is the solution of the system $\dot{V}_p(t) = (1-p)(t-a)^{p-2}x(t)$, with $V_p(a) = 0$ and

$$ \begin{aligned} A(\alpha,N) &= \frac{1}{\Gamma(1-\alpha)}[1+\sum_{p=2}^{N}\frac{\Gamma(p-1+\alpha)}{\Gamma(\alpha)(p-1)!}], \\ B(\alpha,N) &= \frac{1}{\Gamma(2-\alpha)}[1+\sum_{p=1}^{N}\frac{\Gamma(p-1+\alpha)}{\Gamma(\alpha-1)p!}], \\ C(\alpha,p) &= \frac{1}{\Gamma(\alpha-1)\Gamma(2-\alpha)} \times \frac{\Gamma(p-1+\alpha)}{(p-1)!}. \end{aligned} $$

The right Riemann-Liouville fractional derivative is approximated by

$$ {}_tD_b^\alpha x(t) \approx A(\alpha,N)(b-t)^{-\alpha}x(t) - B(\alpha,N)(b-t)^{1-\alpha}\dot{x}(t) + \sum_{p=2}^{N} C(\alpha,p)(b-t)^{1-\alpha-p}W_p(t), $$

where $W_p(t)$ is the solution of the system $\dot{W}_p(t) = -(1-p)(b-t)^{p-2}x(t)$, with $W_p(b) = 0$, and $A(\alpha,N)$, $B(\alpha,N)$ and $C(\alpha,p)$ are as above.

The authors study some particular cases for which restrictions are imposed on the end time T or on $x(T)$.

Numerical methods are discussed by Tricaud and Chen [146]. In this work the fractional differentiation operator used in the OPFDC is approximated using Oustaloup's approximation into a state space realization form, and the OPFDC is reformulated into an integer OCP by using RIOTS-95, a Matlab toolbox to solve this problem. The problem considered in this work is

$$ \text{Minimize} \qquad J(u) = G(x(a),x(b)) + \int_a^T L(t,x(t),u(t))dt, $$

$$ \text{subject to} \qquad {}_aD_t^\alpha x(t) = F(t,x(t),u(t)), $$

with initial condition $x(a) = x_a$ and with the following constraints

$$ u_{\min}(t) \le u(t) \le u_{\max}(t), $$

$$x_{\min}(a) \leq x(a) \leq x_{\max}(a),$$

$$L_{ti}^{v}(t, x(t), u(t)) \leq 0,$$

$$G_{ei}^{v}(x(a), x(b)) \leq 0,$$

$$G_{ee}^{v}(x(a), x(b)) = 0,$$

where $L(\cdot)$, $G(\cdot)$ and $F(\cdot)$ are arbitrary given non-linear functions. The subscripts ti, ei, and ee on the functions $L(\cdot)$, $G(\cdot)$ are trajectory constraint, endpoint inequality constraint and endpoint equality constraint, respectively. The idea is to use the approximation $s^{\alpha} = \prod_{n=1}^{N} \frac{1+s/w_{z,n}}{1+s/w_{p,n}}$ to transform the fractional differentiation operator into an integer operator as follows:

$$_aD_t^{\alpha}x(t) \approx \begin{cases} \dot{z}(t) = & Az(t) + Bu(t), \\ x(t) = & Cz(t) + Du(t), \end{cases}$$

where $A = \begin{bmatrix} -b_{N-1} & -b_{N-2} & \dots & -b_1 & -b_0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$, $D = d$, and

$C = \begin{bmatrix} C_{N-1} & C_{N-2} & \dots & C_1 & C_0 \end{bmatrix}$.

Then, the OPFDC is converted to the following integer OCP:

$$\text{Minimize} \qquad J(u) = G(Cz(a) + Du(a), Cz(b) + Du(b)) + \int_a^b L(t, Cz(t) + Du(t), u(t))dt,$$

$$\text{subject to} \qquad {}_aD_t^{\alpha}x(t) = Az(t) + BF(t, CZ(t) + Du(t), u(t)),$$

with initial condition $z(a) = \frac{x_a w}{Cw}$, where $w = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}^T$, and with the following constraints:

$$u_{\min}(t) \leq u(t) \leq u_{\max}(t),$$

$$x_{\min}(a) \leq Cz(a) + Du(a) \leq x_{\max}(a),$$

$$L_{ti}^{v}(t, Cz(t) + Du(t), u(t)) \leq 0,$$

$$G_{ei}^{v}(Cz(a)+Du(a), Cz(b)+Du(b)) \leq 0,$$

$$G_{ee}^{v}(Cz(a)+Du(a), Cz(b)+Du(b)) = 0.$$

The state $x(t)$ of the initial OPFDC can be retrieved form $x(t) = Cz(t) + Du(t)$. The resulting setting is appropriate as an input for RIOTs-95 Matlab Toolbox.

Kamocki [81] presents optimality conditions of the Pontryagin type for OPFDC under convexity assumptions of the velocity set and cost function, in which the fractional dynamic system involves the Riemann-Liouville derivative. Moreover, he stated the optimality conditions in two cases, first with the initial condition $x_0 = 0$ and secondly when it is different from zero ($x_0 \neq 0$). He consider the following OPFDC:

$$\begin{aligned}
\text{Minimize} \qquad & J[x,u] = \int_a^b f_0(t, x(t), u(t))dt, \\
\text{subject to} \qquad & {}_aD_t^{\alpha} x(t) = f(t, x(t), u(t)), \quad t \in [a,b] \text{ a.e.}, \\
& {}_aI_t^{1-\alpha} x(a) = x_0, \\
& u(t) \in M \subset \mathbb{R}^m, \quad t \in [a,b],
\end{aligned}$$

where $f : [a,b] \times \mathbb{R}^n \times M \to \mathbb{R}^n$, $f_0 : [a,b] \times \mathbb{R}^n \times M \to \mathbb{R}$, $0 < \alpha < 1$, and $x_0 \in \mathbb{R}^n \setminus \{0\}$. This problem is stated under the following assumptions on the data:

(H1) the function $f \in C^1$ with respect to $x \in \mathbb{R}^n$;

(H2) the function $f_0(\cdot, x, u)$ is measurable on $[a,b]$ for all $x \in \mathbb{R}^n$, $u(t) \in M$, and $f_0(t, x, \cdot)$ is continuous on M for a.e. $t \in [a,b]$ and all $x \in \mathbb{R}^n$;

(H3) the function $f_0 \in C^1$ with respect to $x \in \mathbb{R}^n$, for $1 < p < \frac{1}{1-\alpha}$, $\alpha \in (0,1)$, satisfies

$$\begin{aligned}
|f_0(t,x,u)| &\leq a_1(t) + c_1 |x|^p, \\
|(f_0)_x(t,x,u)| &\leq a_2(t) + c_2 |x|^{p-1},
\end{aligned}$$

for a.e. $t \in [a,b]$ and all $x \in \mathbb{R}^n$, where $a_2 \in L^{p'}([a,b], \mathbb{R}_0^+)$, $(\frac{1}{p} + \frac{1}{p'} = 1)$, $a_1 \in L^1([a,b], \mathbb{R}_0^+)$, and $c_1, c_2 \geq 0$;

(H4) the functions $f_x(\cdot, x, u)$, $(f_0)_x(\cdot, x, u)$ are measurable on $[a,b]$ for all $x \in \mathbb{R}^n$, $u \in M$;

(H5) the functions $f_x(t, x, \cdot)$, $(f_0)_x(t, x, \cdot)$ are continuous on M for a.e. $t \in [a,b]$ and all $x \in \mathbb{R}^n$;

(H6) the set M is compact;

(H7) for a.e. $t \in [a,b]$ and all $x \in \mathbb{R}^n$ the set

$$\left\{(f_0(t,x,u), f(t,x,u)) \in \mathbb{R}^{n+1} : \quad u \in M\right\},$$

is convex.

Let $\mathcal{U}_M$ be the set valued mapping defined by

$$\mathcal{U}_M = \left\{u(\cdot) \in L^1([a,b], \mathbb{R}^m) : \quad u(t) \in M, t \in [a,b]\right\}.$$

Therefore, he derived the optimality conditions as follows: if the pair

$$(x^*(\cdot), u^*(\cdot)) \in \left({}_aI_t^\alpha(L^p) + \left\{\frac{d}{(t-a)^{1-\alpha}}, \quad d \in \mathbb{R}^n\right\}\right) \times \mathcal{U}_M,$$

is a locally optimal solution of the problem stated before, then there exists a function $\lambda \in {}_tI_b^\alpha(L^{p'})$ such that

$${}_tD_b^\alpha \lambda(t) = f_x^T(t, x^*(t), u^*(t))\lambda(t) - (f_0)_x(t, x^*(t), u^*(t)) \quad \text{for a.e. } t \in [a,b],$$

$${}_tI_b^{1-\alpha}\lambda(b) = 0,$$

and

$$f_0(t, x^*(t), u^*(t)) - \lambda(t) f(t, x^*(t), u^*(t)) = \min_{u \in M} \left\{f_0(t, x^*(t), u(t)) - \lambda(t) f(t, x^*(t), u(t))\right\},$$

for a.e. $t \in [a,b]$. Furthermore, he derived optimality conditions for an initial condition $x_0 = 0$ (see Kamocki [81]).

Chapter 3

Formulation of Fractional Optimal Control Problems

3.1 Introduction

By FOCPs we denote an OCPs for which either the performance index and/or the dynamical system displays at least one fractional operator.

Although several FOCPs formulations are possible, in general we will consider one for which the performance index is the integral of a function that depends on both the state and the control variables, and the dynamical constraints are described by FDEs. The reason to use FDEs to describe dynamic systems in FOCPs is because we consider instances in which fractional derivatives provide a description of the behavior of the dynamic system which is more accurate than the one given by integer derivatives. More specifically, this is particularly relevant for systems with long-range memory and non-local effects. FOCPs can be defined with respect to different definitions of fractional derivatives. However, the ones in the sense of Riemann-Liouville and of Caputo have been used more widely. There are also a number of specific numerical techniques to solve FOCPs.

FOCPs has a much wider application range of dynamic-control problems with respect to fractional calculus of variations (FCVs), which in its simplest version is defined by minimizing a cost

$$J[x(\cdot)] = \int_a^b L(t, x(t), {}_aD_t^\alpha x(t))dt,$$

subject to boundary conditions

$$x(a) = x_a, \quad x(b) = x_b.$$

For more details on FCVs (see *e.g.*, Agrawal [2], Malinowska and Torres [107], and Malinowska *et al.* [106]).

The main focus of this thesis is FOCPs. Their versatility and wider range of applications constitute key advantages over FCVs. As it will be clear, the methods applied for FOCPs differ substantially from those for FCVs. For example, FOCPs employs the Pontryagin maximum principle, in which the general maximum condition (*i.e.*, $\max_u H(\cdot)$) allows the control variable to be discontinuous (jumping at the boundary point) and this forces the consideration of the space of absolutely continuous functions for the state variable.

3.2 A General Formulation of Fractional Optimal Control Problems

There are several definitions for FOCPs because the diverse types of fractional derivatives make it impossible to consider a typical problem that represents all possibilities. We will consider a FOCP as follows.

We shall be concerned with a given interval $[a, b] \subseteq \mathbb{R}$. We are given a multifunction $\mathcal{U}$ mapping $[a, b] \to \mathbb{R}^m$, and a control is a selection $u(\cdot)$ for $\Omega \subset \mathbb{R}^m$. The function $u(\cdot)$ may be either measurable, continuous, integrable, piecewise continuous or defined otherwise, depending on the problem, satisfying $u(t) \in \Omega$ in a time t, with $a \leq t \leq b$.

We are given a dynamic function $F : [a, b] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$. A fractional trajectory (or state) $x(\cdot)$ on $[a, b]$, corresponding to the control $u(\cdot)$, satisfies the fractional dynamical system (FDS)

$$ {}_aD_t^\alpha x(t) = F(t, x(t), u(t)), $$

where the point $x(a)$ is free to be chosen within a given set C_0. Our aim is to find a control $u(t)$ for FOCP that minimizes the cost function

$$ J[x, u] = \int_a^b L(t, x(t), u(t))dt, $$

where $J[x, u]$ is a performance index (or a cost function), and $L(\cdot)$ is a running cost (or Lagrangian).

In summary, a FOCP is given by

$$ \begin{aligned} (P)\ \text{Minimize} \quad & J[x, u] = \int_a^b L(t, x(t), u(t))dt \\ \text{subject to} \quad & {}_aD_t^\alpha x(t) = F(t, x(t), u(t)) \quad \text{a.e. } t \in [a, b], && (3.1) \\ & u \in \mathcal{U}, && (3.2) \\ & x(a) = x_0 \in \mathbb{R}^n \quad \text{a.e. } t \in [a, b]. && (3.3) \end{aligned} $$

Here, ${}_aD_t^\alpha$ represents a fractional differential operator, such as Riemann-Liouville operator, Caputo operator, Jumarie operator, or another one; $x(\cdot)$ is the state variable, t represents time, and $L : [a, b] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ and $F : [a, b] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ are two given mappings.

Moreover, the FOCPs in which the cost function is given, as in problem (P), are known as problems in the Lagrange form, or Lagrange FOCPs. There are two other FOCPs, the first being the Bolza FOCP, in which there is a terminal cost $g(a, x(a), b, x(b))$ in addition to the running cost in the performance index $J[x, u]$. The second one is Mayer FOCP, in which the performance index only consists in the terminal cost and there is no running cost (*i.e.*, $L(\cdot) = 0$).

A fractional trajectory is a solution of the fractional differential equation (3.1) with the boundary condition (3.3) and for a given control function satisfy (3.2).

Any pair $(x; u)$ satisfies the fractional dynamic is called control process. If a control process whose fractional trajectory remains in the boundary of the attainable set (a set of state space points that can be reached from the initial state with admissible control strategies) is called Boundary process.

Minimum is a solution of FOCPs, and this minimum can be global or local. For instance, consider a function $f : \mathbb{R}^n \to \mathbb{R}$, $D \subseteq \mathbb{R}^n$, then a point $x^* \in D$ is a local minimum of $f(\cdot)$ over D if there exists $\varepsilon > 0$ such that for all $x \in D$ satisfying $|x - x^*| < \varepsilon$, we have

$$f(x^*) \leq f(x), \tag{3.4}$$

i.e., x^* is a local minimum if in some ball around it, $f(\cdot)$ does not attain a value smaller than $f(x^*)$. If (3.4) holds for all $x \in D$, then the minimum is global over D. However, in this thesis, we consider a local minimum of the Pontryagin type, which state that the control Hamiltonian must take a minimum value over feasible controls in the feasible set.

A maximum principle which govern solutions to the problem (P) will be obtained under some assumptions on the data such as:

(H1) The cost function is Lower semi-continuous and Lipschitz,

(H2) $\Omega(t)$ is compact valued set valued map $\forall t \in [a, b]$, and $t \to \Omega(t)$ is $\mathcal{B}-$measurable,

(H3) the function F is continuous,

(H4) the function $F(t, \cdot, u)$ is Lipschitz such that

$$|F(t, x_1, u) - F(t, x_2, u)| \leq k(t, u) |x_1 - x_2| .$$

In this thesis we state, discuss and derive necessary conditions of optimality not only for the FOCP (P) but also from this same problem with additional constraints on the state variable of the form $h(t, x(t)) \leq 0$, for all $t \in [a, b]$.

One important observation is in order. Necessary conditions of optimality are only meaningful if the existence of solution is guaranteed. There are various sets of sufficient conditions for the existence of solution to FOCPs. For instance, if the cost functional is at least lower semicontinuous and, at the same time, the set defined by the constraints of the problem is compact, in a sense defined in a context of compatible topologies, then a solution exists. However, in this thesis, we assume that such an optimal-control process already exists and we are concerned only with the assumptions on the data of the problem under which the necessary conditions of optimality can be derived.

Chapter 4

Maximum Principle for the basic Fractional Optimal Control Problem

4.1 Introduction

In this chapter, we present a new general formulation of FOCPs, we consider FOCP for which the performance index is given by an integral of fractional order, and the dynamics is a mapping specifying the Caputo fractional derivative of the state variable with respect to time. The reasons to choose the Caputo fractional derivative is because it is the most popular one among physicists and scientists, and also the fact that the fractional derivative of constants are zero. Moreover, the assumptions that we impose on the data of the problem enables a novel approach to the proof based on a generalization of Taylor's expansion and a fractional mean value theorem. Another contribution of this chapter consists on an analytic method to solve the fractional differential equation. This is illustrated by an example based on a generalization of the Mittag-Leffler function.

Our approach consists into converting the FOCP into an equivalent OCP and, then, decoding the obtained necessary conditions of optimality into the data of the original problem. It has two key advantages relatively to the alternatives often adopted in the literature: i) more precise insight inherent to the use of variational methods in the original modeling framework; and ii) more direct approximating computational procedures guided by the maximum principle conditions.

It should be remarked that our result differs substantially from the one presented by Kamocki [81] where, by using the quite different calculus of variations approach, necessary conditions of optimality are derived for a different OCP that requires the velocity set (*i.e.*, the set of time derivatives of the state variable) to be convex. This is a very strong assumption and constitutes a key difference from our result which covers dynamic control systems whose velocity sets might be a mere discrete set of points. Moreover, our approach is much more in line with the celebrated classic work of Pontryagin *et al.* [127].

This chapter is organized as follows. In the next Section, we state, discuss, and prove necessary conditions of optimality in the form of a Pontryagin Maximum Principle for non-linear FOCPs. In Section 4.4, a simple illustrative example of a FOCP solved by a method based on the Mittag-Leffler function is presented. Finally, in Section 4.5, we present some conclusions of this research as well as some open challenges. The results of this chapter have been announced in [8]

4.2 The Statement and Assumptions

In this section, we discuss the FOCP considered in this chapter, state the associated necessary conditions of optimality, and present its proof which uses an approach that differs from the ones usually adopted in the literature for this class of FOCPs.

Let us consider the simple general problem as follows

$$\begin{aligned}
(\bar{P})\ \text{Minimize} \quad & {}_{t_0}I^{\alpha}_{t_f} L(t,\bar{x}(t),u(t)) \\
\text{subject to} \quad & {}^{C}_{t_0}D^{\alpha}_{t}\bar{x}(t) = \bar{f}(t,\bar{x}(t),u(t)), \quad [t_0,t_f]\ \mathcal{L}-\text{a.e.}, && (4.1)\\
& \bar{x}(t_0) = \bar{x}_0 \in \mathbb{R}^n, && (4.2)\\
& u(t) \in \mathcal{U}, && (4.3)
\end{aligned}$$

where $\mathcal{U} = \{u : [t_0,t_f] \to \mathbb{R}^m : u(t) \in \Omega(t)\}$, $\Omega : [t_0,t_f] \to \mathbb{R}^m$ is a given set valued mapping, $L : [t_0,t_f] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ and $\bar{f} : [t_0,t_f] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ are given functions defining respectively the running cost (or Lagrangian) functional and the fractional dynamics, ${}_{t_0}I^{\alpha}_{t_f}$ is the Riemann-Liouville fractional integral and ${}^{C}_{t_0}D^{\alpha}_{t}$ is the left Caputo fractional derivative of order $0 < \alpha \leq 1$ of the state variable with respect to time.

It is not hard to see that a simple transformation allows us to convert the problem $(\bar{P})$ into an equivalent one, simply by defining an additional state variable component y by

$${}^{C}_{t_0}D^{\alpha}_{t} y(t) = L(t,\bar{x}(t),u(t)),$$

satisfying the initial condition $y(t_0) = 0$. Then, we conclude that problem $(\bar{P})$ is equivalent to the one as follows

$$\begin{aligned}
(P)\ \text{Minimize} \quad & g(x(t_f)) \\
\text{subject to} \quad & {}^{C}_{t_0}D^{\alpha}_{t}x(t) = f(t,x(t),u(t)), \quad [t_0,t_f]\ \mathcal{L}-\text{a.e.}, && (4.4)\\
& x(t_0) = x_0 \in \mathbb{R}^n, && (4.5)\\
& u(t) \in \mathcal{U}, && (4.6)
\end{aligned}$$

where now $g(x(t_f)) = y(t_f)$, the state variable $x = \begin{bmatrix} y \\ \bar{x} \end{bmatrix}$, *i.e.*, it includes y as a first component with initial value at 0, and the mapping $f = \begin{bmatrix} L \\ \bar{f} \end{bmatrix}$, *i.e.*, it has L as first component.

From now on, we consider this as the basic FOCP in the normal form. We remark that the above problem statement is the simplest one that can be considered while containing all the ingredients required by a "bona fide" OCP.

Now, we will state the assumptions on the data of the problem under which our result will be proved.

(H1) The function g is C_1 in $\mathbb{R}^n$, *i.e.*, continuously differentiable in its domain.

(H2) The function f is C_1 and Lipschitz continuous with constant K_f in x, for all $(t, u) \in \{(t, \Omega(t)) : t \in [t_0, t_f]\}$.

(H3) The function f is continuous in (t, u), for all $x \in \mathbb{R}^n$.

(H4) The set valued map $\Omega : [t_0, t_f] \to \mathbb{R}^m$ is compact valued.

(H5) There is $M > 0$ such that $|f(t, x, \Omega(t))| < M$, for all $(t, x) \in [t_0, t_f] \times \mathbb{R}^n$.

These are, by no means, the weakest hypotheses enabling the proof of the maximum principles for FOCPs. However, they are of interest in that they allows the particularly simple proof adopted in this chapter.

The existence of optimal solutions for the linear case has been recently discussed by Kamocki [80]. Existence of solutions is essential to ensure the meaning of the necessary conditions of optimality. Although this chapter does not concern conditions for the existence of a solution to (P) – as the maximum principle assumes a given optimal-control process a priori – it is not difficult to conclude that this is the case under the assumptions (H1)-(H5).

Indeed, the existence of a solution to (P) is guaranteed if: (i) the cost functional g is at least lower semi-continuous; and (ii) the set of points of the state space, $\mathcal{R}(t_f; t_0, x_0)$, that can be reached at the final time t_f is compact. Observe that, while condition (i) is implied by (H1), condition (ii) follows from (H5) the fact that $t_f < \infty$ which implies that $\mathcal{R}(t_f; t_0, x_0)$ is bounded, and these together with (H2) and (H3) and the convexifying effect of the integration of the dynamics implies that $\mathcal{R}(t_f; t_0, x_0)$ is closed.

Consider

$$H(t, x, p, u) := p^T f(t, x, u),$$

with $p \in \mathbb{R}^n$, to be the Pontryagin function associated to problem (P).

4.3 Maximum Principle of Optimality

Before we are going to the main theorem, we will introduce some definitions which help us to prove this theorem.

Definition 4.1. *Let* $0 < \alpha \leq 1$*. An* α*-Lebesgue point of an integrable function* $f : \mathbb{R} \to \mathbb{R}$ *is a point* $t_0 \in \mathbb{R}$ *satisfying*

$$\lim_{\varepsilon \to 0^+} \frac{1}{2\varepsilon} \, {}_{t_0-\varepsilon}I^{\alpha}_{t_0+\varepsilon} \, |f(t) - f(t_0)|^{2-\alpha} = 0.$$

By directly using the definition α-fractional integral it is easy to conclude that the set of Lebesgue points includes those of the integer order integral (see *e.g.*, Karapetyants and Ginsburg [86]). It is important to point out the well known fact that the subset of Lebesgue points of an integrable function $f(\cdot)$ constitutes a full Lebesgue measure subset (see *e.g.*, Taylor [142]).

In what follows, the definition of fractional state transition matrix (FSTM) for a linear time varying vector valued fractional differential equation of the type

$$ {}^{C}_{a}D^{\alpha}_{t} x(t) = A(t)x(t), \qquad x(a) = x_a, \tag{4.7}$$

where $A(\cdot) : \mathbb{R} \to \mathbb{R}^{n\times n}$, and $x(\cdot) \in \mathbb{R}^n$ is required.

Definition 4.2. *The matrix valued map* $\Phi_\alpha : \mathbb{R} \times \mathbb{R} \to \mathbb{R}^{n\times n}$ *is the FSTM to the linear fractional differential equation (4.7) on the interval* $[a, t]$*, if it satisfies*

$$ {}^{C}_{a}D^{\alpha}_{t} \Phi_\alpha(t,s) = A(t)\Phi_\alpha(t,s), \quad \Phi_\alpha(t,t) = I_n, \ \textit{and } \Phi_\alpha(t,s) = 0_n\,, \ \textit{if } t < s, \tag{4.8}$$

where I_n and 0_n are, respectively, the identity and zero matrices of order n. It is a simple to conclude that $x(t) = \Phi_\alpha(t,a)x_a$ is solution to (4.7).

Definition 4.3. *A pair* $(\bar{x}^*, u^*)$ *is an optimal-control process for problem* $(\bar{P})$ *if it yields a cost lower than that associated with any other feasible control process.*

Obviously, given the equivalence between (P) and $(\bar{P})$ *the same definition holds for* (x, u) *with respect to (P).*

Theorem 4.1. *Let* (x^*, u^*) *be optimal-control process for* (P)*.*

Then, there exists a function $p : [t_0, t_f] \to \mathbb{R}^n$ *satisfying*

- *the adjoint equation*

$$ {}_{t}D^{\alpha}_{t_f} p^T(t) = p^T(t) D_x f(t, x^*(t), u^*(t)), \tag{4.9}$$

where the operator ${}_{t}D^{\alpha}_{t_f}$ *is right Riemann-Liouville fractional derivative, and*

- *the transversality condition*

$$-p^T(t_f) = \nabla_x g(x^*(t_f)), \tag{4.10}$$

- $u^* : [t_0, t_f] \to \mathbb{R}^m$ *is a control strategy such that* $u^*(t)$ *maximizes* $[t_0, t_f]$ $\mathcal{L}$*-a.e., the map*

$$u \to H(t, x^*(t), p(t), u),$$

on $\Omega(t)$*.*

Proof. The first key idea is that any perturbation of the optimal-control $u^*(\cdot)$ that affects the final value of the state trajectory can not strictly decrease the cost. Thus, the proof relies on the comparison between the optimal trajectory $x^*(\cdot)$ and trajectories $x(\cdot)$ which are obtained by perturbing the optimal-control $u^*(\cdot)$.

Let τ be a Lebesgue point in (t_0, t_f), and $\varepsilon > 0$ sufficiently small so that $\tau - \varepsilon \geq t_0$. The Lebesgue point in the fractional context define in the Definition 4.1.

Now, let us consider the perturbed control strategy $u_{\tau,\varepsilon}$ defined by

$$u_{\tau,\varepsilon}(t) = \begin{cases} \bar{u}(t), & \text{if } t \in [\tau - \varepsilon, \tau), \\ u^*(t), & \text{if } t \in [t_0, t_f] \setminus [\tau - \varepsilon, \tau), \end{cases} \tag{4.11}$$

where $\bar{u}(\cdot) \in \Omega(t)$ for all $t \in [\tau - \varepsilon, \tau)$, being τ a Lebesgue point of the reference optimal-control strategy. Note that, there is no loss of generality of the choice of τ due to the fact that the set Lebesgue points is of full Lebesgue measure.

Let $x_{\tau,\varepsilon}(\cdot)$ be the trajectory associated with $u_{\tau,\varepsilon}(\cdot)$, and with $x_{\tau,\varepsilon}(t_0) = x_0$. Clearly, by definition of optimality of (x^*, u^*),

$$\begin{cases} 0 & \leq & g(x_{\tau,\varepsilon}(t_f)) - g(x^*(t_f)) \\ & = & \nabla_x g(x^*(t_f))[x_{\tau,\varepsilon}(t_f) - x^*(t_f)] + \mathrm{o}(\varepsilon), \end{cases} \tag{4.12}$$

where $\nabla_x g(\cdot)$ is the gradient of $g(\cdot)$, and $\mathrm{o}(\varepsilon)$ is a set of function r satisfying $\lim\limits_{\varepsilon \to 0} \dfrac{r(\varepsilon)}{\varepsilon} = 0$. Observe that $x_{\tau,\varepsilon}(t) = x^*(t)$, for all $t \in [t_0, \tau - \varepsilon)$.

Moreover, it is clear that, for all $t \in [\tau - \varepsilon, \tau)$, we have,

$$\begin{aligned} |x_{\tau,\varepsilon}(t) - x^*(t)| &\leq {}_{\tau-\varepsilon}I_\tau^\alpha |f(s, x_{\tau,\varepsilon}(s), \bar{u}(s)) - f(s, x^*(s), u^*(s))| \\ &\leq {}_{\tau-\varepsilon}I_\tau^\alpha K_f |x_{\tau,\varepsilon}(s) - x^*(s)| + 2M \frac{\varepsilon^\alpha}{\Gamma(\alpha+1)} \\ &\leq \frac{\overline{M}\varepsilon^\alpha}{\Gamma(\alpha+1)}, \end{aligned} \tag{4.13}$$

where

$$\overline{M} = 2M\left(1 + K_f \sum_{n=1}^{\infty} \frac{\Gamma(\alpha)^{n-1}}{\Gamma(n\alpha+1)} \varepsilon^{n\alpha}\right).$$

It is not difficult to show that this series converges and thus $\overline{M}$ is some finite positive number. The last inequality was obtained by the next theorem, and, in particular, holds for $t = \tau$.

Theorem 4.2. *Generalized Bellman-Gronwall inequality* (see, *e.g.*, Lin [95]).

Suppose $\alpha > 0$, $t \in [0, T)$ and the functions $a(t)$, $b(t)$ and $w(t)$ are a non-negative and continuous functions on $0 \le t < T$ with

$$w(t) \le a(t) + b(t) \int_0^t (t-s)^{\alpha-1} w(s) ds,$$

where $b(t)$ is a bounded and monotonic increasing function on $[0, T)$. Then,

$$w(t) \le a(t) + \int_0^t \left[\sum_{n=1}^{\infty} \frac{(b(t)\Gamma(\alpha))^n}{\Gamma(n\alpha)} (t-s)^{n\alpha-1} a(s)\right] ds, \quad t \in [0, T).$$

For the proof and more details about Generalized Bellman-Gronwall inequality (see Appendix B).

Before proceeding with the proof, we need the following auxiliary result.

In what follows, let $\Phi_\alpha(\cdot, \cdot)$ denote the state transition matrix (see Definition 4.2) for the linear fractional differential system

$${}_{t_0}^{C}D_t^{\alpha} \xi(t) = D_x f(t, x^*(t), u^*(t)) \xi(t).$$

Lemma 4.1. *Consider the time interval $[a, b]$ and the function $F(t, x(t)) = f(t, x(t), u(t))$, where $u(t)$ is a general feasible control function. Moreover, consider $\tilde{x}(\cdot)$, $y(\cdot)$, and $\tilde{x}_\nu(\cdot)$ to be, respectively, the solutions to the following fractional differential systems define on the interval $[a, b]$:*

- ${}_{a}^{C}D_t^{\alpha} \tilde{x}(t) = F(t, \tilde{x}(t))$ *with* $\tilde{x}(a) = x_a$,
- ${}_{a}^{C}D_t^{\alpha} y(t) = D_x F(t, \tilde{x}(t)) y(t)$ *with* $y(a) = \bar{y}$, *and*
- ${}_{a}^{C}D_t^{\alpha} \tilde{x}_\nu(t) = F(t, \tilde{x}_\nu(t))$ *with* $\tilde{x}_\nu(a) \in x_a + \nu^\alpha \bar{y} + \mathrm{o}(\nu^\alpha) \mathrm{B}_1^{\mathrm{n}}(0)$.

Then, for all ν positive and sufficiently small real number, we have that $\tilde{x}_\nu(\cdot)$ satisfies on the time interval $[a, b]$

$$\tilde{x}_\nu(t) \in \tilde{x}(t) + \nu^\alpha y(t) + \mathrm{o}(\nu^\alpha) \mathrm{B}_1^{\mathrm{n}}(0).$$

Here, $B_1^n(0)$ denotes the closed unit ball of $\mathbb{R}^n$ centered at 0.

Proof. Let us consider the first-order expansion of the map $x \to F(t, \cdot)$ around $\tilde{x}(t)$. We have

$$F(t, \tilde{x}_\nu(t)) - F(t, \tilde{x}(t)) - D_x F(t, \tilde{x}(t))(\tilde{x}_\nu(t) - \tilde{x}(t)) \in \mathrm{o}(\|\tilde{\mathrm{x}}_\nu(\mathrm{t}) - \tilde{\mathrm{x}}(\mathrm{t})\|),$$

for all $t \in [a, b]$. Since

$$ {}_a^C D_t^\alpha y(t) = D_x F(t, \tilde{x}(t)) y(t),$$

with $\nu^\alpha y(a) \in \tilde{x}_\nu(a) - \tilde{x}(a) + \mathrm{o}(\nu^\alpha)\mathrm{B}_1^{\mathrm{n}}(0)$, we have that

$$ {}_a^C D_t^\alpha [\tilde{x}_\nu(t) - \tilde{x}(t) - \nu^\alpha \, y(t)] = \zeta(t),$$

for some $\zeta \in L^1$ satisfying $\zeta(t) \in \mathrm{o}(\nu^\alpha)\mathrm{B}_1^{\mathrm{n}}(0)$ in L^1.

By integrating, we have

$$\tilde{x}_\nu(t) - \tilde{x}(t) - \nu^\alpha y(t) = \tfrac{1}{\Gamma(\alpha)} \int_a^t (t-\tau)^{\alpha-1} (F(\tau, \tilde{x}_\nu(\tau)) - F(\tau, \tilde{x}(\tau))$$

$$-\nu^\alpha D_x F(\tau, \tilde{x}(\tau)) y(\tau)) d\tau + \tfrac{1}{\Gamma(\alpha)} \int_a^t (t-\tau)^{\alpha-1} \zeta(\tau) d\tau.$$

Before continuing, we should observe that it is a simple exercise to conclude that the assumption $(H2)$ implies that $\|D_x F(t, x(t))\| \leq K_f$. Note also that it is not difficult to see that $\beta(t) := \frac{1}{\Gamma(\alpha)} \int_a^t (t-\tau)^{\alpha-1} \zeta(\tau) d\tau \in \mathrm{o}(\nu^\alpha)$, for all $t \in [a, b]$,

Now, by putting $z(t) = \tilde{x}_\nu(t) - \tilde{x}(t) - \nu^\alpha y(t)$, using the above observations and the assumption $(H2)$, we obtain the inequality

$$\|z(t)\| \leq \beta(t) + \frac{K_f}{\Gamma(\alpha)} \int_a^t (t-\tau)^{\alpha-1} \|z(\tau)\| d\tau.$$

Theorem 4.2 (Generalized Bellman-Gronwall inequality) yields

$$\|z(t)\| \leq \|\beta(t)\| + \int_a^t \left[\sum_{n=1}^{\infty} \frac{K_f^n \Gamma(\alpha)^n}{\Gamma(n\alpha)} (t-\tau)^{n\alpha-1} (\tau - a) \|\beta(\tau)\| \right] d\tau.$$

Let

$$\bar{\beta} = \sup_{t \in [a,b]} \|\beta(t)\|.$$

Obviously, we have that $\bar{\beta} \in \mathrm{o}(\nu^\alpha)$. By performing the integral in the right hand side of the above inequality, we have that, for all $t \in [a, b]$,

$$\|z(t)\| \leq \bar{\beta} \left(1 + \sum_{n=1}^{\infty} \frac{K_f^n \Gamma(\alpha)^n}{\Gamma(n\alpha + 1)} (t-a)^{n\alpha} \right),$$

which, by using the one parameter Mittag-Leffler function (see Chapter 2) can be expressed by

$$\|z(t)\| \leq \bar{\beta} E_{\alpha,\alpha}(K_f \Gamma(\alpha)(t-a)^{\alpha}).$$

Thus, Lemma 4.1 is proved, *i.e.*, for all $t \in [a, b]$,

$$\tilde{x}_{\nu}(t) \in \tilde{x}(t) + \nu^{\alpha} y(t) + \mathrm{o}(\nu^{\alpha}) \mathrm{B}_1^{\mathrm{n}}(0). \tag{4.14}$$

□

Now, by considering (4.12) and applying Lemma 4.1 on the time interval $[\tau, t_f]$ with $F(t, x) = f(t, x, u_{\tau,\varepsilon}(t))$, $\nu^{\alpha} = \varepsilon$, $a = \tau$, $t = t_f$, $\tilde{x}_{\nu} = x_{\tau,\varepsilon}$ and $\tilde{x} = x^*$, we conclude that

$$\begin{cases} 0 & \leq & \nabla_x g(x^*(t_f))[x_{\tau,\varepsilon}(t_f) - x^*(t_f)] + \mathrm{o}(\varepsilon) \\ & \leq & \varepsilon \nabla_x g(x^*(t_f)) y(t_f) + \mathrm{o}(\varepsilon) \\ & = & \varepsilon \nabla_x g(x^*(t_f)) \Phi_{\alpha}(t_f, \tau) y(\tau) + \mathrm{o}(\varepsilon), \end{cases} \tag{4.15}$$

where $\Phi_{\alpha}(t_f, \tau)$ is the fractional state transition matrix associated with the linear system

$${}_a^C D_t^{\alpha} y(t) = D_x F(t, x^*(t)) y(t),$$

in the interval $[\tau, t_f]$. By letting

$$-p^T(t_f) = \nabla_x g(x^*(t_f)),$$

and

$$p^T(t) = p^T(t_f) \Phi_{\alpha}(t_f, t), \tag{4.16}$$

we conclude that the adjoint variable $p(\cdot)$ satisfies the right Riemann-Liouville fractional linear equation

$${}_t D_{t_f}^{\alpha} p^T(t) = p^T(t) D_x F(t, x^*(t)),$$

i.e., $p(\cdot)$ satisfies the adjoint equation of our maximum principle as well as the associated transversality condition.

Finally, by putting together (4.15), and (4.16) and by choosing

$$y(\tau) = f(\tau, x^*(\tau), \bar{u}) - f(\tau, x^*(\tau), u^*(\tau)),$$

were $\bar{u} = \bar{u}(\tau)$, we obtain

$$0 \geq \varepsilon p^T(\tau)[f(\tau, x^*(\tau), \bar{u}) - f(\tau, x^*(\tau), u^*(\tau))] + \mathrm{o}(\varepsilon).$$

By dividing both sides of this inequality by $\varepsilon > 0$ and by taking the limit $\varepsilon \to 0^+$, we conclude the inequality

$$0 \geq p^T(\tau)[f(\tau, x^*(\tau), \bar{u}) - f(\tau, x^*(\tau), u^*(\tau))],$$

which, from the arbitrariness of $\bar{u} \in \Omega(t)$, yields the maximum condition at time $t = \tau$,

$$H(\tau, x^*(\tau), p(\tau), u^*(\tau)) \geq H(\tau, x^*(\tau), p(\tau), \bar{u}).$$

The fact that τ is an arbitrary Lebesgue point in $[t_0, t_f]$ implies that the maximum condition of our main result holds, that is, $u^*(t)$ maximizes on $\Omega(t)$, the map $u \to H(t, x^*(t), p(t), u)$, $[t_0, t_f]$ $\mathcal{L}$-a.e..

Our main result is proved. □

4.4 Illustrative Example

The Pontryagin maximum principle proved in the previous section is now apply to solve a simple problem of resources management that involves minimizing a certain fractional integral subject to given controlled FDEs.

We consider the following problem

$$\begin{aligned}
&\text{Minimize} && J(u) && (4.17)\\
&\text{subject to} && {}_0^C D_t^\alpha x(t) = u(t)x(t), \quad t \in [0, T], && (4.18)\\
& && x(0) = x_0, && (4.19)\\
& && u(t) \in [0, 1], && (4.20)
\end{aligned}$$

where $J(u) = {}_0I_T^\alpha (1 - u(t))x(t)$, with $0 < \alpha < 1$ and $T > \Gamma(\alpha + 1)^{\alpha^{-1}}$. Here, ${}_0I_T^\alpha$ is fractional integral and ${}_0^C D_t^\alpha$ is left Caputo fractional derivative.

The variable x represents a natural resource that takes positive values (note that $x_0 > 0$ necessarily) "grows" according to the law (4.18), where the function $u(\cdot)$, designated by control, represents the fraction of the available resource that is used to promote further growth.

The overall goal is to find the control strategy that maximizes the amount of accumulated resource over the time interval $[0, T]$ given by the fractional integral (4.17).

First, we consider an additional state variable component y, satisfying

$${}_0^C D_t^\alpha y(t) = (1 - u(t))x(t), \quad y(0) = 0,$$

in order obtain the canonic problem statement in the form considered in our main result, that is

$$\begin{array}{ll} \text{Minimize} & y(T) \\ \text{subject to} & {}_0^C D_t^\alpha x(t) = u(t)x(t), \qquad x(0) = x_0, \\ & {}_0^C D_t^\alpha y(t) = (1-u(t))x(t), \quad y(0) = 0, \\ & u(t) \in [0,1]. \end{array}$$

From Theorem 4.1, the adjoint equation (4.9) and the transversality condition (4.10) for this problem are

$$ {}_t D_T^\alpha p_1(t) = [p_1 u^*(t) + p_2(1-u^*(t))], \qquad p_1(T) = 0, \tag{4.21}$$

$$ {}_t D_T^\alpha p_2(t) = 0, \qquad p_2(T) = 1, \tag{4.22}$$

where ${}_t D_T^\alpha$ is right Riemann-Liouville fractional derivative of order α. Thus, we have that $p_2(t) \equiv p_2(T) = 1$, and equation (4.21) becomes

$$ {}_t D_T^\alpha p_1(t) = [(p_1(t) - 1)u^*(t) + 1]. \tag{4.23}$$

From the maximum condition, we know that $u^*(t)$ maximizes, $\mathcal{L}$-a.e. in $[0,1]$, the mapping

$$v \to p^T(t) f(t, x^*(t), y^*(t), v) = [p_1(t)v + p_2(t)(1-v)]x^*(t).$$

Since $p_2(t) = 1$ and $x^*(t) > 0$ for all $t \in [0,T]$ (this is concluded from the fact that $x_0 > 0$), the mapping to be maximized can be simplified to $v \to (p_1(t) - 1)v$. Thus, given that the system is time invariant, we have that

$$u^*(t) = \begin{cases} 1, & \text{if } p_1(t) > 1, \\ 0, & \text{if } p_1(t) < 1. \end{cases}$$

Since $p_1(T) = 0$, and $p_1(\cdot)$ is continuous, $\exists\, b > 0$ s.t. $u^*(t) = 0$ $\forall t \in [T-b, T]$. Thus, from (4.21), we have ${}_t D_T^\alpha p_1(t) = 1$ and, by backwards integration we obtain

$$p_1(t) = \frac{(T-t)^\alpha}{\Gamma(\alpha+1)}. \tag{4.24}$$

Obviously that, for $t^* = T - (\Gamma(\alpha+1))^{\frac{1}{\alpha}}$, we obtain $p_1(t^*) = 1$. Now, Let us determine the optimal-control for $t < t^*$. Since, independently of the control $p_1(\cdot)$ remains monotonically decreasing, we have for $t < t^*$, $u^*(t) = 1$, and, thus,

$$ {}_t D_{t^*}^\alpha p_1(t) = p_1(t). \tag{4.25}$$

The solution of this linear fractional differential equation (4.25) is given by

$$p_1(t) = p_1(t^*)\Phi_\alpha(t^*, t),$$

where $p_1(t^*) = 1$ and $\Phi_\alpha(t^*, t)$ is the FSTM (in fact, scalar-valued) that can be computed by the Mittag-Leffler function defined in Chapter 2.

By setting $\beta = \alpha$, $A = [1]$ and by replacing t by $t^* - t = T - \Gamma(\alpha+1)^{\alpha^{-1}} - t$, we conclude that

$$p_1(t) = (t^* - t)^{\alpha-1} \sum_{k=0}^{\infty} \frac{(t^* - t)^{k\alpha}}{\Gamma((k+1)\alpha)}.$$

Note that if $\alpha = 1$, then we have classical solution e^{T-t-1}. Since we have the optimal-control u^*, we can easily compute the optimal trajectory which satisfies $x^*(0) = x_0$, and

$$ {}_0^C D_t^\alpha x^*(t) = \begin{cases} x^*(t), & \text{if } t \in [0, t^*], \\ 0, & \text{if } t \in [t^*, T]. \end{cases}$$

We can compute the optimal trajectory $x^*(t)$ by the generalization Mittag-Leffler function. For all $t \in [0, t^*]$, $x^*(0) = x_0$, we conclude that

$$\begin{aligned} x^*(t) &= E_\alpha(at^\alpha) \\ &= x_0 E_\alpha(t^\alpha) \\ &= x_0 \sum_{k=0}^{\infty} \frac{t^{k\alpha}}{\Gamma(k\alpha+1)}. \end{aligned}$$

Note that if $\alpha = 1$, then we have classical solution $x_0 e^t$.

Now, we compute the optimal trajectory $x^*(t)$ in the interval $[t^*, T]$, which $u^*(t) = 0$, $x^*(t) = x^*(T)$, we conclude that

$$\begin{aligned} x^*(t) &= x^*(t^*) \\ &= x_0 E_\alpha((t^*)^\alpha) \\ &= x_0 \sum_{k=0}^{\infty} \frac{(T - (\Gamma(\alpha+1))^{\frac{1}{\alpha}})^{k\alpha}}{\Gamma(k\alpha+1)}. \end{aligned}$$

whose optimum cost is

$$y^*(T) = x_0 E_\alpha((T)^\alpha).$$

Note that if $\alpha = 1$, then we have classical solution $x_0 e^{T-1}$.

4.5 Conclusion

This chapter concerns the derivation of necessary conditions of optimality in the form of Pontryagin maximum principle for a non-linear FOCP whose differential equation involves the Caputo derivative of the state variable with respect to time. Under mild assumptions on the data of the problem the proof involved the direct application of variational arguments, thus avoiding the often used argument of converting the OCP into a conventional one and, then, express the optimality conditions for this auxiliary problem back in the fractional derivative context. Another interesting novelty consists in the fact that, unlike in most FOCP formulations, we consider the cost functional given by a fractional integral of Riemann-Liouville type.

A simple example illustrating the application of our maximum principle was presented. The optimal-control strategy was computed analytically being the fractional differential adjoint equation solved by using technique based on a generalization Mittag-Leffler function.

A natural sequel of this chapter concerns the weakening of the assumptions on the data of the problem. Notably the mere measurability dependence of the dynamics with respect to time and to the control variables. This will certainly require more sophisticated variational arguments and the use of methods and results of nonsmooth analysis. Another direction of research consists in increasing the structure of the FOCP by considering additional state endpoint constraints, and state and/or mixed constraints in its formulation. In this case, additional regularity assumptions will be needed to ensure that the obtained necessary conditions of optimality do not degenerate.

Chapter 5

Fractional Integration and Measure Concepts

In this chapter, we will introduce some results that will be of relevance in the next chapters. Here, we focus our attention on the definition of fractional integral with respect to general Radon measures, that is, the fractional Stieltjes integral, due to the fact that these measures appear as multipliers associated with state constraints in the maximum principle for FOCPs. We adopt the Jumarie fractional integral operator and its properties to achieve our purposes. The prerequisites for understanding our main results are presented in the first section of this chapter, as these concepts are important to develop the main results.

5.1 Introduction

In this section, we present some basic concepts of the theory of measure and integration required to develop the necessary auxiliary results.

5.1.1 Measures

The measure theory is intimately connected with integration. Specifically, the concept of measure generalizes the concept of length of an interval, area of a rectangle, volume of a parallelepiped, etc. The theory of measure and integration can be found in a number of books (see *e.g.*, Carter and Van Brunt [38], Folland [58], Natanson [114], Taylor [142], Thomson [143], Thomson *et al.* [144], and Yeh *et al.* [154]).

The families of sets that serve us the domains of measures, such as algebra, $\sigma-$algebra and Borel $\sigma-$algebra, are defined in Appendix D.

Let X be a set equipped with a $\sigma-$algebra Ω. A set function $\mu(\cdot)$ is called a measure on Ω (or on (X, Ω)) if it satisfies the following conditions:

(i) $\mu(\emptyset) = 0$, where $\emptyset$ is a null set;

(ii) if $\{S_j\} \subset \Omega$ is a countable collection of disjoint sets, then

$$\mu\left(\bigcup_j S_j\right) = \sum_j \mu(S_j).$$

The pair (X, Ω) is called a measurable space, and the sets in Ω are called measurable sets. Moreover, if $\mu(\cdot)$ is a measure on (X, Ω), then a triple (X, Ω, μ) is called a measure space. Some of the basic properties of measures are summarized in Appendix D.

A measure $\mu\colon \mathcal{B}(X) \to [0, \infty]$ is called a Borel measure if $\mu(K) < \infty$ for every compact set $K \subset \mathcal{B}(X)$, where $\mathcal{B}(X)$ is the Borel $\sigma-$algebra in X. Some measures, such as regular measure, outer regular measure, inner regular measure and Radon measure, are defined in Appendix D.

Definition 5.1. *(Mutual singular measures) Two measures μ and ν, defined on a $\sigma-$algebra of X, are called mutual singular measures if there are disjoint sets A and B such that (i) $X = A \cup B$, (ii) μ is zero in all measurable subsets of A, and (iii) ν is zero in all measurable subsets of B.*

For the particular case where ν is the Lebesgue measure on X, we simply say that the measure μ is singular.

Definition 5.2. *(Absolutely continuous measures) A measure ν on a measurable space (X, Ω) is absolutely continuous with respect to the measure μ on the same measurable space (we write $\nu \ll \mu$), if for any measurable set $E \in \Omega$, we have*

$$\mu(E) = 0 \Rightarrow \nu(E) = 0.$$

Definition 5.3. *(Atomic measure) An atomic measure, or a discrete measure, is a measure that only takes non-zero values on discrete subsets.*

A set $A \in \Omega$ is called an atom of μ, if (i) $\mu(A) > 0$, and (ii) for every $V \in \Omega$, with $V \subset A$, one has $\mu(V) = 0$ or $\mu(V) = \mu(A)$.

A regular Borel measure can be written as in the canonical decomposition of the sum of an absolutely continuous measure (μ_{ac}), a singular continuous measure (μ_{sc}) and an atomic measure (μ_a), that is,

$$\mu = \mu_{ac} + \mu_{sc} + \mu_a.$$

5.1.2 Integration

The classical definition of an integral was firstly proposed by Cauchy and later developed by Riemann, who defined the integral $\int_a^b f(x)dx$ as a limit of the so-called Riemann sums

(Natanson [114]). (Details on Riemann integrals are presented in Appendix D.) The functions for which the Riemann integral exists are said to be integrable in the Riemann sense. In order for the function $f(\cdot)$ to be Riemann-integrable, it is necessary that it should be bounded and that it does not exhibit any point of accumulation of non-summable discontinuities. The proof of this statement can be found in many books (see *e.g.*, Carter and Van Brunt [38], Folland [58], Taylor [142], and Yeh *et al.* [154]). Thus, this significant restrictivity inherent to the definition of Riemann integral motivates the introduction of the Lebesgue integral.

The Lebesgue integral is a generalization of the Riemann integral. For instance, to study the Riemann integral one needs to subdivide the interval of integration into a finite number of subintervals, but in the Lebesgue integral the interval is subdivided into more general sets called measurable sets. On the other hand, the Lebesgue integral makes no distinction between bounded and unbounded set in integration. There are numerous convincing arguments for considering the Lebesgue integration (see *e.g.*, Carter and Van Brunt [38], De Barra [50], Folland [58], Natanson [114], Taylor [142], and Yeh *et al.* [154]). Details on Lebesgue integral for simple function, bounded measurable functions, and non-negative functions are exposed in Appendix D.

Now, we are going to present the definition of a further generalization of integrals to the so-called the Stieltjes integral.

Definition 5.4. *Let $\psi(\cdot)$ be a function defined on the interval $t \in [a, b]$, and $\mu(\cdot)$ be a Radon measure. Let us partition the interval $[a, b]$ by n points $a < t_1 < t_2 < \cdots < t_n < b$, and put $t_0 = a$, $t_{n+1} = b$, $\Delta t_i = t_{i+1} - t_i$, $i = 0, 1, \ldots, n$, and*

$$D_i = \begin{cases} [t_i, t_{i+1}[, & i = 0, \cdots, n-1, \\ \\ [t_i, t_{i+1}], & i = n. \end{cases}$$

By choosing a point ξ_i in each D_i, we have that the limit

$$\lim_{\max \Delta t_i \to 0} \sum_{i=0}^{n} \psi(\xi_i)\mu(D_i),$$

is called the Stieltjes integral of the function $\psi(\cdot)$ with respect to the measure $\mu(\cdot)$, and is denoted by $\int_{[a,b]} \psi(t) d\mu(t)$. Here, $\mu(D_i) = \mu(t_{i+1}^-) - \mu(t_i^-)$, for $i = 0, \cdots, n-1$, and $\mu(D_n) = \mu(b^+) - \mu(t_n^-)$.

By Radon measure, it is meant an inner regular Borel measure see Definition D.4 in Appendix D. (For more details and important properties of the Stieltjes integral, see *e.g.*, Bray [34], Carter and Van Brunt [38], Natanson [114], and Thomson [143]).

As mentioned before, our focus is on fractional integrals, among which the Riemann-Liouville one is the most common. In order to investigate our challenge, we will exploit the relation between the Riemann-Liouville and Jumarie fractional integrals, already stated in Chapter 2. The next definition displays the Riemann-Liouville fractional integral for a function with respect to another function.

Definition 5.5. *Let $h\colon [a,b] \to \mathbb{R}$ be a monotone increasing function on $(a,b]$, and h' its continuous derivative on (a,b). The fractional integral of order $\alpha > 0$ of a function $f(\cdot)$ with respect to the function $h(\cdot)$ on $[a,b]$, is defined as*

$$_aI^{\alpha}_{t;h} f(t) = \frac{1}{\Gamma(\alpha)} \int_a^t (h(t) - h(\tau))^{\alpha-1} f(\tau) h'(\tau) d\tau.$$

Note that, if $h(t) = t$, $t \in [a,b]$, the fractional integral $_aI^{\alpha}_{t;h} f(t)$ will be the usual Riemann-Liouville fractional integration (Samko *et al.* [135]).

Now, we pursue with the key purposes of this chapter.

5.2 Fractional Integration with respect to a Measure

In this section, we define the integral in a fractional context with respect to a general Radon measure, and introduce it in the Jumarie fractional integral form.

Definition 5.6. *Let $f(\cdot)$ be a function defined on the interval $t \in [a,b]$, and $\mu(\cdot)$ be a Radon measure. Let us partition the interval $[a,b]$ by n points $a < t_1 < t_2 < \cdots < t_n < b$, and put $t_0 = a$, $t_{n+1} = b$, $\Delta t_i = t_{i+1} - t_i$, $i = 0, 1, \ldots, n$, and*

$$D_i = \begin{cases} [t_i, t_{i+1}[, & i = 0, \cdots, n-1, \\ \\ [t_i, t_{i+1}], & i = n. \end{cases}$$

By choosing a point ξ_i in each D_i, we call to the limit

$$\frac{1}{\Gamma(\alpha+1)} \lim_{\max \Delta t_i \to 0} \sum_{i=0}^{n} f(\xi_i)(\mu(D_i))^{\alpha}, \tag{5.1}$$

the fractional Stieltjes integral of order α of the function $f(\cdot)$ with respect to the measure $\mu(\cdot)$, and denoted by

$$\frac{1}{\Gamma(\alpha+1)} \int_{[a,b]} f(t)(d\mu(t))^{\alpha},$$

where $\alpha > 0$. If $\alpha = 1$, we have the classic Stieltjes integral which as stated in the Definition 5.4.

Proposition 5.1. *Let $\mu(\cdot)$ be a positive Borel measure on $\mathbb{R}$. Then, the fractional integral for the function $f\colon \mathbb{R} \to [0,\infty]$ with respect to the measure $\mu(\cdot)$ is given by*

$$
{}_aJ^{\alpha}_{t;\mu} f(t) = \begin{cases} \frac{1}{\Gamma(\alpha+1)} \int_{[0,t]} f(\bar{t})(d\mu(\bar{t}))^{\alpha}, & \textit{if } t \in \mathrm{Supp}(\mu_c), \\ \\ \frac{1}{\Gamma(\alpha+1)} \int_{[0,t)} f(\bar{t})(d\mu(\bar{t}))^{\alpha} + V, & \textit{if } t \in \mathrm{Supp}(\mu_a), \end{cases} \tag{5.2}
$$

where $V = \frac{f(t)(\mu(\{t\}))^{\alpha}}{\Gamma(\alpha+1)}$, $\mu(\{t\}) = \mu(t^+) - \mu(t^-)$, μ_a is atomic measure, $\mu_c = \mu_{ac} + \mu_{sc}$ is atomless measure, μ_{ac} and μ_{sc} are defined before, and ${}_aJ^{\alpha}_{t;\mu}$ is operator of Jumarie fractional integral with respect to the measure $\mu(\cdot)$.

Proof. The idea of the proof is using some change of variables to transform the integer Stieltjes integral to Lebesgue integral. Then, the usual expression of the fractional integral as a convolution given by an integer integral is written down in terms of the original parametrization.

Two cases have to be considered: the first one when the measure atomless *i.e.*, $t \in \mathrm{Supp}(\mu_c)$, and when we have atomic measure *i.e.*, $t \in \mathrm{Supp}(\mu_a)$.

When $t \in \mathrm{Supp}(\mu_c)$, $\mu_c = \mu_{ac} + \mu_{sc}$, the integer Stieltjes integral defined by

$$
\int_{[0,t]} f(\bar{t}) d\mu(\bar{t}). \tag{5.3}
$$

By using the time change of variables, in which for $t \in [0,T]$, there is $s \in [0, \mu([0,T])]$ satisfying

$$
s(t) = \int_{[0,t]} d\mu(\bar{t}),\ (i.e.,\ ds = d\mu(t)).
$$

Moreover, for all integrable functions $f : [0,T] \to \mathbb{R}^n$, $t = \theta(s)$ and $s = \sigma(t)$, exists $\tilde{f} : [0, \sigma(T)] \to \mathbb{R}^n$ such that

$$
\tilde{f}(s) = (f \circ \theta)(s),
$$

consequently, there exist a Lebesgue integral that is equivalent to the Stieltjes integral that mentioned in (5.3) such that

$$
\int_{[0,t]} f(\bar{t}) d\mu(\bar{t}) = \int_{[0,s]} \tilde{f}(\bar{s}) d\bar{s}. \tag{5.4}
$$

Now, we consider the fractional integral of $\tilde{f}(\cdot)$ for $\alpha \in (0,1]$, which can be expressed in the reparametrized time variable as follows

$$
\frac{1}{\Gamma(\alpha)} \int_0^s \tilde{f}(\bar{s})(s - \bar{s})^{\alpha-1} d\bar{s}. \tag{5.5}
$$

By using the assumptions stated before, then (5.5) satisfies

$$\frac{1}{\Gamma(\alpha)} \int_{[0,s]} \tilde{f}(\sigma(\bar{t}))(\sigma(t) - \sigma(\bar{t}))^{\alpha-1} d\sigma(\bar{t}). \tag{5.6}$$

Since $t > \bar{t}$, we have that

$$\mu(t) - \mu(\bar{t}) = \begin{cases} \mu([\bar{t}, t)), & \text{if } t < T, \\ \\ \mu([\bar{t}, T]), & \text{if } t = T. \end{cases}$$

Thus, by using a change of variables in (5.6), we conclude that

$$\frac{1}{\Gamma(\alpha)} \int_{[0,t]} f(\bar{t}) \left(\mu(t) - \mu(\bar{t})\right)^{\alpha-1} d\mu(\bar{t}).$$

This equation represent Riemann-Liouville fractional integral with respect to measure $\mu(\cdot)$ for the case in which $t \in \mathrm{Supp}(\mu_c)$, by applying the relation between Riemann-Liouville and Jumarie fractional integral operator as in Chapter 2, where the integrating measure $d\mu(\bar{t})$ replaces $d\tau$. Consequently the fractional integral with respect to the measure $d\mu(\cdot)$ in Jumarie fractional integral form is, for $t \in \mathrm{Supp}(\mu_c)$, defined by

$$\frac{1}{\Gamma(\alpha+1)} \int_{[0,t]} f(\bar{t})(d\mu(\bar{t}))^{\alpha}.$$

When $t \in \mathrm{Supp}(\mu_a)$, the function $f(\cdot)$ have a discontinuity at a point $t \in [t^-, t^+]$, then we have

$$\frac{1}{\Gamma(\alpha)} \int_{\mu(t^-)}^{\mu(t^+)} (\mu(t^+) - s)^{\alpha-1} f(t) m(s) ds,$$

where $m(s)$ is the measure distribution function of the measure μ satisfy

$$m(s) = \begin{cases} 1, & \text{if } s \in [\mu(t^-), \mu(t^+)], \\ 0, & \mathit{otherwise}. \end{cases}$$

Consequently, we conclude that

$$\frac{f(t)}{\Gamma(\alpha+1)} (\mu(t^+) - \mu(t^-))^{\alpha}.$$

So, we can be write this equation as follows

$$\frac{f(t)}{\Gamma(\alpha+1)} \left(\mu(\{t\})\right)^{\alpha},$$

where $\mu(\{t\}) = \mu(t^+) - \mu(t^-)$. Thus, the general case when $t \in \mathrm{Supp}(\mu_a)$ is given by

$$\frac{1}{\Gamma(\alpha+1)}\int_{[0,t)} f(\bar{t})(d\mu(\bar{t}))^{\alpha}+\frac{1}{\Gamma(\alpha+1)}f(t)(\mu(\{t\})^{\alpha}.$$

□

Remark 5.1. *When* $t \in \mathrm{Supp}(\mu_c) \cup \mathrm{Supp}(\mu_a)$, *we can write the measure as the canonical decomposition of atomic and non-atomic measure,* i.e. $\mu = \mu_c + \mu_a$, *as follows*

$$\frac{1}{\Gamma(\alpha+1)}\int_{[0,t]} f(\bar{t})(d\mu_c(\bar{t}))^{\alpha}+\sum_{t_i\le t}\frac{f(t_i)}{\Gamma(\alpha+1)}\left(\mu(\{t_i\})\right)^{\alpha}.$$

5.3 Fundamental Properties

Here, we introduce some properties for the fractional Stieltjes integral, which are well known in the literature for Lebesgue fractional integral calculus in Jumarie context, such as the change of variable. A standard formula of change of variable of fractional integral introduced by Jumarie [75, 77]

$$\int f(x)(dx)^{\alpha} = \int f(g(t))(g'(t))^{\alpha}(dt)^{\alpha}, \quad \alpha \in (0,1),$$

where we made the substitution $x = g(t)$ in which $g(t)$ be a non-decreasing differential function. We derive such a formula for fractional Stieltjes integral as follows: first we consider a strictly increasing function $g\colon \mathbb{R} \to \mathbb{R}$ on the interval I, *i.e.*, if for all $t_1, t_2 \in I$ with $t_1 < t_2$, then, $g(t_1) < g(t_2)$. Furthermore, let the function $g(\cdot)$ be continuous and strictly increasing on the interval I, then we can say that $g(I)$ be an interval defined by

$$g(I) = \{g(t)\colon t \in I\}.$$

Proposition 5.2. *(Fractional Stieltjes change of variable). Let I be any interval and $g\colon \mathbb{R} \to \mathbb{R}$ a continuous function strictly increasing on the interval I. Let $\mu\colon \mathbb{R} \to \mathbb{R}$ be a Radon measure. Then,*

$$\int_I (f\circ g)(s)\,(d(\mu\circ g)(s))^{\alpha} = \int_{g(I)} f(t)(d\mu(t))^{\alpha},$$

where $(f\circ g)(s)$ denotes the composition of $f(s)$, and $g(s)$ is defined by $(f\circ g)(s) = f\left[g(s)\right]$, for all $s \in I$.

Proof. For the sake of simplify, we assume that $t = g(s)$, $\tilde{\mu}(s) = (\mu\circ g)(s)$, $\tilde{f}(s) = (f\circ g)(s)$ and $ds = (g^{-1})'(t)dt$.

As we know, the measure have a canonical decomposition such that, for $\mu_c = \mu_{ac} + \mu_{sc}$, we have $\mu = \mu_c + \mu_a$ satisfying

$$d\mu = d\mu_c + d\mu_a.$$

Therefore,

$$d\tilde{\mu} = d\tilde{\mu}_c + d\tilde{\mu}_a.$$

By definition of $\tilde{\mu}(\cdot)$, we have for all $A \subset I$

$$\int_A d\tilde{\mu}(s) = \int_{g(A)\cap S_c} d\mu_c(t) + \sum_{t_i \in g(A)\cap S_a} \mu_a(\{t_i\}),$$

where $S_c = S_{ac} + S_{sc}$ and S_a are the supports of the measures μ_c and μ_a, respectively. Let B be the support of the measure μ such that $\mu(B) = \|\mu\|_{TV}$, and $\mu(B^c) = 0$.

According to the results of Section 5.2, for $s \in [0, S]$, and $t \in [0, T]$, we have

$$\begin{aligned} \sigma(s) &= \int_{[0,s]} (d\tilde{\mu}_c(s) + d\tilde{\mu}_a(s)) \\ &= \int_0^s d\tilde{\mu}_c(s) + \sum_{s_i \in [0,s]} \tilde{\mu}_a(\{s_i\}). \end{aligned}$$

Consequently,

$$d\sigma(s) = d\tilde{\mu}_c(s) + \tilde{\mu}_a(s)\tilde{\delta}_s,$$

where $\tilde{\delta}_s$ is a Dirac delta impulse. Furthermore, for $s_1 \leq S$ we conclude that

$$\frac{1}{\Gamma(\alpha+1)} \int_{[0,s_1]} \tilde{f}(s)(d\tilde{\mu}(s))^\alpha = \frac{1}{\Gamma(\alpha)} \int_{[0,s_1]} \tilde{f}(s)(\sigma(s_1) - \sigma(s))^{\alpha-1} d\sigma(s).$$

By using the definition of $\sigma(\cdot)$. We have

$$\frac{1}{\Gamma(\alpha+1)} \int_{[0,s_1]} \tilde{f}(s)(d\tilde{\mu}(s))^\alpha = \frac{1}{\Gamma(\alpha)} \int_{[0,s_1]} \tilde{f}(s) \left(\int_s^{s_1} d\tilde{\mu}_c(s) + \sum_{s_i \in [s,s_1]} \tilde{\mu}_a(\{s_i\}) \right)^{\alpha-1} d\sigma(s) \quad (5.7)$$

where we have used the assumptions $t_i = g(s_i)$, $f(t) = (\tilde{f} \circ g^{-1})(t)$, $\mu_a(\{t_i\}) = \tilde{\mu}_a(\{g^{-1}(t_i)\})$, and $\mu_c(t) = (\tilde{\mu}_c \circ g^{-1})(t)$ such that $d\mu_c(t) = d\left((\tilde{\mu}_c \circ g^{-1})(t)\right)(g^{-1})'(t)$. Then, the right-hand side of (5.7) satisfies

$$\frac{1}{\Gamma(\alpha)} \int_{[0,t_1]} f(t) \left(\int_t^{t_1} d\mu_c(t) + \sum_{t_i \in [t,t_1]} \mu_a(\{t_i\}) \right)^{\alpha-1} d\mu(t), \quad (5.8)$$

where $d\mu(t) = d\mu_c(t) + \mu_a(t)\delta_t$. Therefore, (5.8) will be

$$\frac{1}{\Gamma(\alpha)} \int_{[0,t_1]} f(t) \left(\mu([t, t_1])\right)^{\alpha-1} d\mu(t) = \frac{1}{\Gamma(\alpha+1)} \int_{[0,t_1]} f(t) \left(d\mu(t)\right)^\alpha.$$

Proposition 5.2 are proved. □

Chapter 6

Necessary Conditions of Optimality for Fractional Nonsmooth Differential Inclusion Problems with State Constraints

6.1 Introduction

Differential inclusions of fractional order have recently been addressed by several researcher for many problems and a several results related to fractional differential inclusions have appeared in the literature (see *e.g.*, Ahmad and Ntouyas [7], Benchohra *et al.* [30], Benchohra *et al.* [31], Cernea [39], El-Sayed and Ibrahim [56], and Kamocki and Obczyński [83]).

The main contribution of this chapter is the formulation of necessary conditions in the form of a Maximum Principle for nonsmooth optimal-control problem with state constraints and with dynamics given by fractional differential inclusion. Here, we consider the Jumarie fractional derivative

$$x^{(\alpha)}(t) \in F(t, x(t)), \qquad \mathcal{L} - \text{a.e.}, \tag{6.1}$$

where $x^{(\alpha)}(\cdot)$ is fractional Jumarie derivative of order $\alpha \in (0,1)$, and $F(t,x(t))$ is a set valued map (multifunction) defined on $[a,b] \times \mathbb{R}^n$.

Consider the following fractional differential inclusion problem

$$
(P) \qquad \begin{cases} \text{Minimize} & g(x(b)) \\ \text{subject to} & x^{(\alpha)}(t) \in F(t, x(t)), \quad t \in [a, b], \\ & h(t, x(t)) \leq 0, \quad a \leq t \leq b, \\ & x(a) \in C, \end{cases}
$$

where the function $g : \mathbb{R}^n \to \mathbb{R}$ is a given cost function , $x^{(\alpha)}$ is the fractional Jumarie derivative of order $\alpha \in (0, 1]$, $F(t, x)$ is a given multifunction defined on $[a, b] \times \mathbb{R}^n$, the function $h : [a, b] \times \mathbb{R}^n \to \mathbb{R}$ defines the given inequality state constraint, $x(a)$ are initial point, and C is a given subset of $\mathbb{R}^n$.

A feasible trajectory of the problem (P) is a solution of fractional differential inclusion satisfying all the constraints on the problem (P). We say that x^* be a local minimum of the problem (P), if it minimizes the objective function over all other feasible states $x \in \mathbb{R}^n$, in some neighborhood such that $|x - x^*| \leq \varepsilon$.

Definition 6.1. *Let $x \colon [a, b] \to \mathbb{R}^n$ be a given continuous function. We say that this function lies in $\Omega \subset [a, b] \times \mathbb{R}^n$, if the point $(t, x(t))$ is in Ω, for each $t \in [a, b]$. Let $\varepsilon > 0$ be a small positive constant. Then, $\varepsilon -$ tube about x is defined by*

$$
T(x; \varepsilon) = \{(t, \bar{x}) \in [a, b] \times \mathbb{R}^n : a \leq t \leq b, |\bar{x} - x(t)| \leq \varepsilon\}.
$$

$t-$section of $\Omega \subset [a, b] \times \mathbb{R}^n$ defined by

$$
\Omega_t = \{x \in \mathbb{R}^n \colon (t, x) \in \Omega\}, \qquad \forall t \in [a, b].
$$

The multifunction $F(t, \cdot)$ is called Lipschitz of rank $k(t)$ for all x_1 and x_2, if for all $\eta_1 \in \ F(t, x_1)$ there exists $\eta_2 \in F(t, x_2)$ such that

$$
|\eta_1 - \eta_2| \leq k(t) \left| x_1 - x_2 \right|.
$$

The multifunction $F(\cdot, x)$ is Lebesgue measurable if for all open set C in $\mathbb{R}^n$ the set

$$
\{t \in [a, b] : \quad F(t, x) \cap C \neq 0\},
$$

is Lebesgue measurable for all $x \in \mathbb{R}^n$. The measurability of F can be defined equivalently if the set C is an arbitrary closed set (see *e.g.*, Clarke [43], and Vinter [150]). Moreover, a multifunction $F(\cdot, \cdot)$ is measurably Lipschitz on $\Omega \subset [a, b] \times \mathbb{R}^n$, if (i) for each $x \in \mathbb{R}^n$, the multifunction $F(\cdot, x)$ is measurable on $[a, b]$, and (ii) for each $t \in [a, b]$ the multifunction $F(t, \cdot)$ is nonempty and Lipschitz of rank $k(t)$ on Ω_t.

Basic Hypotheses

(H1) The function $g(\cdot)$ is Lipschitz on Ω_b of rank K_g, such that

$$|g(x_1(b)) - g(x_2(b))| \leq K_g |x_1(b) - x_2(b)|.$$

(H2) $F(\cdot, x)$ is closed, and convex valued on Ω.

(H3) F is $\alpha-$integrably bounded on Ω, such that there is an $\alpha-$integrable function $\phi\colon [a, b] \to \mathbb{R}$, such that any measurable selection $\eta(t)$ for a multifunction $F(t, x)$ satisfies

$$|\eta(t)| \leq \phi(t).$$

(H4) The multifunction $(t, x) \to F(t, x)$ is measurably Lipschitz along $\Omega \subset [a, b] \times \mathbb{R}^n$, that is, for each $x \in \mathbb{R}^n$, a multifunction $F(\cdot, x)$ is measurable in $t \in [a, b]$, and for each $t \in [a, b]$, a multifunction $F(t, \cdot)$ is Lipschitz of rank K.

(H5) The function h is upper semicontinuous, and there exists a constant $K_h > 0$ such that the function $h(t, \cdot)$ is Lipschitz of rank K_h on Ω for all $t \in [a, b]$, i.e., it satisfies,

$$|h(t, x_1) - h(t, x_2)| \leq K_h |x_1 - x_2|.$$

Let the Hamiltonian H for the problem (P) is a function $H\colon \Omega \times \mathbb{R}^n \to \mathbb{R}$ be defined by

$$H(t, x, p) = \max\{\langle p, v\rangle : v \in F(t, x)\},$$

This Chapter is organized as follows. In the next Section we introduce notations, definitions, and preliminary facts will be used later in this chapter. In Section 6.3, the necessary optimal conditions are stated and proved. These conditions are used in an essential way in the proof of the Maximum Principle for nonsmooth fractional optimal-control problems addressed in Chapter 7.

6.2 Auxiliary Technical Results

In this section, we present and discuss key results on existence solutions for fractional differential inclusions as well as on the compactness of the set of trajectories for this class of dynamical systems. The basic idea consists in adapting the results proved in Clarke [43] from the integer to the fractional context.

The supremum norm $\|\cdot\|$ is defined by

$$\|x\| := \max_{t \in [a,b]} \{|x(t)|\},$$

where $|\cdot|$ is Euclidean norm.

At this point, some basic definitions and properties on the dynamics are stated. Let $v(\cdot) : [a, b] \to \mathbb{R}^n$ be a Lebesgue measurable selection of the set-valued map $F(\cdot, x(\cdot))$, that is

$$v(t) \in F(t, x(t)), \quad \mathcal{L} - \text{a.a.}$$

The solution $x(t)$ of fractional differential inclusion (6.1) satisfies the fractional differential equation

$$x^{(\alpha)}(t) = v(t),$$

for some $v(\tau)$ Lebesgue measurable selection of $F(\tau, x(\tau))$, and integrable with respect to $(t-\tau)^{\alpha-1}d\tau$, for $t \in [a, b]$, such that $x(\cdot)$ satisfying

$$x(t) = x(a) + \frac{1}{\Gamma(\alpha+1)} \int_a^t v(\tau)(d\tau)^\alpha,$$

is well defined. Here, $\int_a^t (\cdot)(dt)^\alpha$ is Jumarie fractional integral operator (see Chapter 2). Without loss of generality, we consider $\alpha \in (0, 1)$.

Definition 6.2. *Let $d_{F(\cdot,\cdot)}(\cdot) : \Omega \times \mathbb{R}^n \to [0, \infty]$ be the distance function associated to a multifunction $F : \mathbb{R}^m \to \mathbb{R}^n$, defined by*

$$d_{F(t,x)}(v) := \inf_{z \in F(t,x)} \{|v - z|\},$$

and $x(\cdot)$ is an $\alpha-$absolutely continuous function from $[a, b] \to \mathbb{R}^n$ lying in Ω. By definition

$$\boldsymbol{d}^\alpha(x, F) = \frac{1}{\Gamma(\alpha+1)} \int_a^b d_{F(t,x(t))}(x^{(\alpha)}(t))(dt)^\alpha.$$

Obviously, $d_{F(t,x)}(v) = 0$, if and only, if $v \in F(t, x)$.

Proposition 6.1. *Consider the multifunction $F : [a, b] \times \mathbb{R}^n \to \mathbb{R}^n$. Then, for a fixed $t \in [a, b]$, the distance function $d_{F(t,x)}(v)$ is Lipschitz of rank $k(t)$ such that*

$$\left|d_{F(t,x_1)}(v_1) - d_{F(t,x_2)}(v_2)\right| \le k(t) \left|x_1 - x_2\right| + \left|v_1 - v_2\right|, \tag{6.2}$$

where the map $t \to d_{F(t,x)}(v)$ is Lebesgue measurable for any $x \in \mathbb{R}^n$ and any $v \in \mathbb{R}^n$.

Proof. We can write the left-hand side of (6.2) as follows

$$\left|d_{F(t,x_1)}(v_1) - d_{F(t,x_2)}(v_2)\right| \le \left|d_{F(t,x_1)}(v_1) - d_{F(t,x_2)}(v_1)\right| + \left|d_{F(t,x_2)}(v_1) - d_{F(t,x_2)}(v_2)\right|. \tag{6.3}$$

Since $F(t, \cdot)$ is $k(\cdot)$ Lipschitz such that

$$F(t, x_1) \subset F(t, x_2) + k(t) \left|x_1 - x_2\right| B_1(0), \quad \forall x_1, x_2 \in \mathbb{R}^n,$$

we conclude that

$$d_{F(t,x_1)}(v_1) \leq d_{F(t,x_2)}(v_1) + k(t)\,|x_1 - x_2|\,, \quad \forall v_1 \in F(t,x_1).$$

It follows from this relation, and by exchanging the roles of x_1 and x_2, that

$$\left|d_{F(t,x_1)}(v_1) - d_{F(t,x_2)}(v_1)\right| \leq k(t)\,|x_1 - x_2|\,. \tag{6.4}$$

Let $\varepsilon > 0$ be arbitrarily small. Then, from the definition of $d_{F(t,\cdot)}(\cdot)$, there is $z_2 \in F(t,x_2)$ such that

$$\begin{aligned} d_{F(t,x_2)}(v_1) &\geq |v_1 - z_2| - \varepsilon \\ &\geq |v_1 - z_2 - v_2 + v_2| - \varepsilon \\ &\geq |v_2 - z_2| - |v_1 - v_2| - \varepsilon \\ &\geq d_{F(t,x_2)}(v_2) - |v_1 - v_2| - \varepsilon. \end{aligned}$$

Since ε is arbitrarily small, and by exchanging the roles of v_1 and v_2, we conclude that

$$\left|d_{F(t,x_2)}(v_1) - d_{F(t,x_2)}(v_2)\right| \leq |v_1 - v_2|\,. \tag{6.5}$$

Finally, substituting (6.4) and (6.5) in the right-hand side of (6.3), we get the inequality (6.2). Now, we show the measurability of $t \to d_{F(t,x)}(v)$. Let arbitrary open set C_k defined by $C_k = z^* + \varepsilon_k B$, where B is an open unit ball centered at zero. Since F is $\mathcal{L}-$measurable, then the set

$$\{t \in [a,b] : C_k \cap F(t,x) \neq \emptyset\},$$

is $\mathcal{L}-$measurable. Define $W_k(z,t) = \{|v - z| : z \in C_k \cap F(t,x)\}$. It is clear that, for each z, $W_k(z,\cdot)$ is $\mathcal{L}-$measurable. By defining $g_k(t) = \max_z\{W_k(z,t)\}$, we also may assert that $g_k(\cdot)$ is $\mathcal{L}-$measurable. So, we conclude that

$$d_{F(t,x)}(v) = \lim_{k\to\infty} g_k(t),$$

is $\mathcal{L}-$measurable. $\square$

Theorem 6.1. *Let $x(\cdot)$ be an $\alpha-$absolutely continuous function in the $\varepsilon-$tube $T(x;\varepsilon) \subseteq \Omega$, for some constant $\varepsilon > 0$, assume that the multifunction $F\colon \Omega \to \mathbb{R}^n$ is Lipschitz of rank $k(t)$ on Ω, and $\mathbf{d}^\alpha(x,F) < \varepsilon/K$, where $K = E_\alpha\left(\frac{1}{\Gamma(\alpha+1)}\int_a^b k(t)(dt)^\alpha\right)$. Then, there exists a fractional trajectory $y(\cdot)$ for the multifunction F lying in the tube $T(x;\varepsilon)$ satisfying $y(a) = x(a)$, and*

$$\|x - y\| \leq \frac{1}{\Gamma(\alpha+1)}\int_a^b \left|x^{(\alpha)}(t) - y^{(\alpha)}(t)\right|(dt)^\alpha \leq K\mathbf{d}^\alpha(x,F) < \varepsilon.$$

Proof. The idea of the proof is to construct a sequence of approximative fractional absolutely continuous functions $\{x_n(t)\}$, beginning with $x_0(t) \equiv x(t)$, and by choosing $x_{n+1}^{(\alpha)}(t)$ as the closest point to $x_n^{(\alpha)}(t)$ of the set $F(t, x_n(t))$, that will converge for a $\alpha-$trajectory for F. Let $v_n(t)$ be a measurable selection of a multifunction $F(t, x_n(t))$ a.e. such that

$$\left|v_n(t) - x_n^{(\alpha)}(t)\right| = d_{F(t,x_n(t))}(x_n^{(\alpha)}(t)) \text{ a.e..} \tag{6.6}$$

Let $v_0(t) \in F(t, x(t))$ a.e. be such that

$$\left|v_0(t) - x^{(\alpha)}(t)\right| = d_{F(t,x(t))}(x^{(\alpha)}(t)) \text{ a.e..}$$

From the above, $v_0(t)$ is $\alpha-$integrable and, thus,

$$x_1(t) = x(a) + \frac{1}{\Gamma(\alpha+1)} \int_a^t v_0(\tau)(d\tau)^\alpha.$$

is well defined. Consequently, we have

$$x_1^{(\alpha)}(t) = v_0(t).$$

Therefore,

$$\left|x_1^{(\alpha)}(t) - x^{(\alpha)}(t)\right| = d_{F(t,x(t))}(x^{(\alpha)}(t)). \tag{6.7}$$

By applying the fractional Jumarie integral operator as well as the standard modulus integral inequality on both sides of (6.7), we get

$$|x_1(t) - x(t)| \leq \frac{1}{\Gamma(\alpha+1)} \int_a^b d_{F(t,x(t))}(x^{(\alpha)}(t))(dt)^\alpha.$$

From Definition 6.2 and a condition of this result, we conclude that

$$|x_1(t) - x(t)| \leq \mathbf{d}^\alpha(x, F) < \varepsilon/K. \tag{6.8}$$

Thus, $x_1(t)$ is in the tube $T(x; \varepsilon)$ and, therefore, we may choose $v_1(t) \in F(t, x_1(t))$ a.e. such that

$$\left|v_1(t) - x_1^{(\alpha)}(t)\right| = d_{F(t,x_1(t))}(x_1^{(\alpha)}(t)) \text{ a.e..}$$

As before $v_1(t)$ is $\alpha-$integrable and, we may define x_2 by

$$x_2(t) = x(a) + \frac{1}{\Gamma(\alpha+1)} \int_a^t v_1(\tau)(d\tau)^\alpha.$$

Then, we have

$$x_2^{(\alpha)}(t) = v_1(t).$$

i.e.,

$$\left|x_2^{(\alpha)}(t) - x_1^{(\alpha)}(t)\right| = d_{F(t,x_1(t))}(x_1^{(\alpha)}(t)). \tag{6.9}$$

From the Lipschitz condition, we have

$$d_{F(t,x_1(t))}(x_1^{(\alpha)}(t)) \le d_{F(t,x(t))}(x_1^{(\alpha)}(t)) + k(t)\,|x_1(t) - x(t)|\,.$$

Obviously, from above $x_1^{(\alpha)}(t) \in F(t, x(t))$, therefore $d_{F(t,x(t))}\left(x_1^{(\alpha)}(t)\right) = 0$, and, thus, we conclude that

$$\left|x_2^{(\alpha)}(t) - x_1^{(\alpha)}(t)\right| \le k(t)\,|x_1(t) - x(t)|\,. \tag{6.10}$$

By integrating both sides in (6.10) and using (6.8), we conclude that

$$|x_2(t) - x_1(t)| \le \mathbf{d}^{\alpha}(x, F)\frac{1}{\Gamma(\alpha+1)}\int_a^t k(\tau)(d\tau)^{\alpha}. \tag{6.11}$$

Note that we can write

$$\begin{aligned}
|x_2(t) - x(t)| &\le |x_2(t) - x_1(t)| + |x_1(t) - x(t)| \\
&\le \mathbf{d}^{\alpha}(x,F)\tfrac{1}{\Gamma(\alpha+1)}\textstyle\int_a^t k(\tau)(d\tau)^{\alpha} + \mathbf{d}^{\alpha}(x,F) \\
&\le \mathbf{d}^{\alpha}(x,F)\left[\tfrac{1}{\Gamma(\alpha+1)}\textstyle\int_a^t k(\tau)(d\tau)^{\alpha} + 1\right] \qquad (6.12)\\
&\le \mathbf{d}^{\alpha}(x,F)E_{\alpha}\left(\tfrac{1}{\Gamma(\alpha+1)}\textstyle\int_a^t k(\tau)(d\tau)^{\alpha}\right) \\
&\le K\mathbf{d}^{\alpha}(x,F) < \varepsilon,
\end{aligned}$$

where the second and third inequalities are due to (6.8) and (6.11). Thus x_2 lies in $T(x;\varepsilon)$. These two first steps clearly show how the induction process can be constructed. Because of the fact that the successive velocities lie in $T(x;\varepsilon)$, we may, generally, choose $v_n(t) \in F(t, x_n(t))$ a.e. satisfying (6.6), where $v_n(t)$ is $\alpha-$integrable satisfying

$$x_{n+1}^{(\alpha)}(t) = v_n(t).$$

Let

$$\begin{aligned}
M_n^{\alpha}(t) &= {}_aJ_t^{\alpha}\left(k(t_1)\,{}_aJ_{t_1}^{\alpha}\left(k(t_2)\cdots\left({}_aJ_{t_{n-1}}^{\alpha}k(t_n)\right)\right)\right) \\
&= \tfrac{1}{n!}\left(\tfrac{1}{\Gamma(\alpha+1)}\textstyle\int_a^t k(\tau)(d\tau)^{\alpha}\right)^n,
\end{aligned} \tag{6.13}$$

where ${}_aJ_{t_{(\cdot)}}^{\alpha}(\cdot)$ is Jumarie fractional integral operator (see Chapter 2).

Consequently, $x_{n+1}^{(\alpha)}(t) \in F(t, x_n(t))$ satisfies

$$\left|x_{n+1}^{(\alpha)}(t) - x_n^{(\alpha)}(t)\right| \le k(t)\,|x_n(t) - x_{n-1}(t)|\,.$$

Thus, the mathematical induction leads to the recursive relation

$$\left|x_{n+1}^{(\alpha)}(t) - x_n^{(\alpha)}(t)\right| \leq \mathbf{d}^\alpha(x,F)k(t)M_{n-1}^\alpha(t), \quad n = 1,2,\cdots. \tag{6.14}$$

Therefore,

$$|x_{n+1}(t) - x_n(t)| \leq \mathbf{d}^\alpha(x,F)M_n^\alpha(t), \quad n = 0,1,2,\cdots. \tag{6.15}$$

At each step, $x_n(t)$ is in the tube $T(x;\varepsilon)$. It follows that

$$\begin{aligned} |x_n(t) - x(t)| &\leq \mathbf{d}^\alpha(x,F)\textstyle\sum_{j=0}^{n-1}\frac{(M_n^\alpha(b))^j}{j} \\ &\leq \mathbf{d}^\alpha(x,F)E_\alpha\left(\frac{1}{\Gamma(\alpha+1)}\int_a^b k(\tau)(d\tau)^\alpha\right) \qquad (6.16) \\ &\leq \mathbf{d}^\alpha(x,F)K < \varepsilon. \end{aligned}$$

From (6.14), the sequence $\{x_n^{(\alpha)}(t)\}$ is a Cauchy sequence in $L^\alpha([a,b];\mathbb{R}^n)$. Let $v(\cdot) \in L^\alpha([a,b];\mathbb{R}^n)$ be a weak limit of this sequence. From (6.15), we deduce that $\{x_n\}$ convergences to a continuous function y with $y(a) = x(a)$. It follows that

$$v(t) \in F(t,y(t)) \quad \text{a.e.}.$$

So, from

$$x_n(t) = x(a) + \frac{1}{\Gamma(\alpha+1)}\int_a^t x_n^{(\alpha)}(\tau)(d\tau)^\alpha, \quad n = 1,2,\cdots,$$

we conclude that

$$y(t) = x(a) + \frac{1}{\Gamma(\alpha+1)}\int_a^t v(\tau)(d\tau)^\alpha.$$

Therefore,

$$y^{(\alpha)}(t) = v(t).$$

Then, the fractional trajectory $y(t)$ for the multifunction F satisfies

$$\begin{aligned} \|y(t) - x(t)\| &\leq E_\alpha\left(\frac{1}{\Gamma(\alpha+1)}\int_a^b k(t)(dt)^\alpha\right)\mathbf{d}^\alpha(x,F) \\ &\leq K\mathbf{d}^\alpha(x,F). \end{aligned}$$

The result is proved. □

Theorem 6.2. *Let $\Omega \subset [a,b]\times\mathbb{R}^n$ and a multifunction $F\colon \Omega \to \mathbb{R}^n$ be given. Assume that F is $\mathcal{L}\times\mathcal{B}$ measurable with non empty closed convex values on Ω and that F be such that $|F(\cdot,x(\cdot)| \leq \phi(\cdot)$ where $\phi(\cdot)$ is an essentially bounded and $\alpha-$integrable function. Assume also, there is a multifunction $G\colon [a,b] \to \mathbb{R}^n$ and a positive valued function $\pi(t)$ such that the following hypotheses are satisfied*

(1) For all $t \in [a, b]$, $G(t) + \pi(t)B \subset \Omega_t$,

(2) for each $t \in [a, b]$, $x \in G(t) + \pi(t)B$, the multifunction $F(t, \cdot)$ is upper semiconscious,

(3) for each (t, x) in the interior of Ω, the multifunction $F(\cdot, x)$ is measurable.

Let $\{x_i\}$ be a sequence of $\alpha-$differentiable functions on $[a, b]$, $\{r_i\}$ be a sequence of measurable function on $[a, b]$ converging to 0, as $i \to +\infty$, and $\{\Lambda_i\}$ is a sequence of measurable subset of $[a, b]$ such that $\mathcal{L} - meas(\Lambda_i) \to (b - a)$ as $i \to \infty$. Furthermore, suppose that

(4) $\{x_i(t)\} \in G(t)$, and $\left|\{x_i^{(\alpha)}(t)\}\right| \leq \phi(t)$, $t \in [a, b]$ a.e.,

(5) the sequence $\{x_i(a)\}$ is bounded,

(6) $x_i^{(\alpha)}(t) \in F(t, x_i(t)) + r_i(t)B$, $\forall t \in \Lambda_i$ a.e..

Then, there is a subsequence of $\{x_i\}$ converging to an $\alpha-$differentiable function x satisfying Lebesgue a.e.

$$x^{(\alpha)}(t) \in F(t, x(t)).$$

Proof. This theorem concerns the compactness (in some sense) of the set of feasible trajectories as solutions to the $\alpha-$differential inclusion. For the case $\alpha = 1$, we have, from the Arzela-Ascoli theorem, that for a given Banach space, a subset is compact if and only if any equicontinuous and uniformly bounded sequence has a subsequence converging to an element of the set. However, in general, this is not the case if $\alpha \in (0, 1)$. As it follows from Theorem A of (Okada and Ricker [117]), it is clear that, in general, I^α is not a compact operator and thus the Dunford-Pettis property, i.e., all weakly compact operators transform weakly compact sets from a Banach space into norm-compact sets of another Banach space (complete continuity)– in particular, for integer integral operators mapping the space of L^1 integrable functions into that of continuous functions –, cannot be applied. For such values of α, weaker notions of operator compactness are investigated in (Salem and Cichoń [134]). For this purpose, the weaker notion of Pettis integrability is exploited in this article. The function $f : T \to X$, being T a measurable space endowed with the structure (T, Σ, μ), is Pettis integrable over $A \in \Sigma$ if $\exists$ a vector $e \in X$ so that

$$\langle \psi, e \rangle = \int_A \langle \psi, f(t) \rangle \, d\mu(t)$$

for all functionals $\psi \in X^*$, where X^* is the dual of X. From Theorem 8 in (Salem and Cichoń [134]), it follows that $f : T \to X$ is Pettis integrable, then $I^\alpha f$ is defined a.e. in T, f is fractionally integrable on T, and if, additionally, f is strongly measurable (*i.e.*, f

is a.e. equal to the limit of a sequence of measurable countably-valued functions) with $T = [0,1]$, then $I^\alpha f : T \to X$ is bounded, weakly continuous and

$$\sup_{\|\phi\|\leq 1} \left\{ \int_0^1 \phi I^\alpha f(t) dt \right\} \leq \frac{1}{\Gamma(\alpha+1)} \sup_{\|\phi\|\leq 1} \left\{ \int_0^1 \phi f(t) dt \right\}.$$

Moreover, it follows from an easy extension of the remarks of the above theorem that, for $\alpha \leq 1$, that $I^\alpha : L^p([0,1];\mathbb{R}^n) \to L^p([0,1];\mathbb{R}^n)$ is compact, and if $p > \max\{1, \frac{1}{\alpha}\}$, that $I^\alpha : L^p([0,1];\mathbb{R}^n) \to C([0,1];\mathbb{R}^n)$ is also compact. This means that, by the Arzela-Ascoli Theorem that we can choose a sequence of bounded and equicontinuous $\alpha-$differentiable functions $\{x_i\}$ with a subsequence converging to some continuous function x and that, by applying Pettis criterion, we can consider a further subsequence for which the sequence $\{x_i^{(\alpha)}\}$ converges weakly to a limit $\gamma \in L^\alpha([a,b];\mathbb{R}^n)$.

From the hypothesis (5), we know that $\{x_i(a)\}$ is a bounded sequence, and, thus, there is a subsequence of $\{x_i(a)\}$ (we do not relabel that converges to $x(a)$). For such a subsequence, we have

$$x_i(t) = x_i(a) + \frac{1}{\Gamma(\alpha+1)} \int_a^t x_i^{(\alpha)}(\tau)(d\tau)^\alpha,$$

and, thus, we deduce that

$$x(t) = x(a) + \frac{1}{\Gamma(\alpha+1)} \int_a^t \gamma(\tau)(d\tau)^\alpha,$$

thus $x(\cdot)$ is fractional trajectory such that $x^{(\alpha)}(t) = \gamma(t)$ a.e..

Consider an arbitrary Lebesgue measurable set $M \subset [a,b]$ and the Hamiltonian $H(t,z,p)$ defined by

$$H(t,z,p) = \max\left\{\langle p, v\rangle : v \in F(t,z)\right\}.$$

From its definition, we have the for all $t \in M$,

$$v \in F(t,z) \Rightarrow H(t,z,p) \geq \langle p, v\rangle, \qquad \forall p \in \mathbb{R}^n.$$

Consequently, the function $z(\cdot)$ is an $\alpha-$trajectory for F if and only if

$$H(t,z,p) \geq \left\langle p, z^{(\alpha)}\right\rangle, \quad \forall p \in \mathbb{R}^n, \quad \forall t \in M.$$

Thus, from hypotheses (6), it follows that

$$\int_{M\cap\Lambda_i} H(t, x_i(t), p) dt \geq \int_{M\cap\Lambda_i} \left\langle p, x_i^{(\alpha)}(t)\right\rangle dt - \int_{M\cap\Lambda_i} r_i(t)\,|p|\, dt - \bar{\phi}(b-a-|\Lambda_i|),$$

where $|\Lambda_i|$ is the Lebesgue measure of the set Λ_i and $\bar{\phi}$ is the essential supremum of the function $\phi(\cdot)$. By taking upper limit, as $i \to +\infty$, we have

$$\limsup_{i\to\infty} \int_{M\cap\Lambda_i} H(t, x_i(t), p)dt \geq \limsup_{i\to\infty} \int_{M\cap\Lambda_i} \left\langle p, x_i^{(\alpha)}(t) \right\rangle dt$$

$$- \limsup_{i\to\infty} \int_{M\cap\Lambda_i} (r_i(t)|p| + \bar{\phi}(b - a - |\Lambda_i|))dt.$$

From the hypotheses, we have that the elements of the sequence $\{x_i^{(\alpha)}\}$ are $\alpha-$integrably bounded, $\|r_i\| \to 0$, and $\mathcal{L} - \text{meas}(\Lambda_i) \to (b - a)$. Then, we conclude that

$$\int_M \limsup_{i\to\infty} H(t, x_i(t), p)dt - \limsup_{i\to\infty} \int_M \left\langle p, x_i^{(\alpha)}(t) \right\rangle dt \geq 0.$$

From the upper semicontinuity of H, we have

$$\limsup_{i\to\infty} H(t, x_i(t), p) \leq H(t, x(t), p).$$

Then, we conclude that

$$\int_M \left(H(t, x(t), p) - \left\langle p, x^{(\alpha)}(t) \right\rangle \right) dt \geq 0.$$

Since in the previous inequality M is an arbitrary Lebesgue measurable set, we conclude that

$$H(t, x(t), p) \geq \left\langle p, x^{(\alpha)}(t) \right\rangle, \quad t \in [a, b] \text{ a.e..}$$

Since H is continuous in p, this inequality can be obtained $\forall p \in \mathbb{R}^n$.

Thus it follow that

$$x^{(\alpha)}(t) \in F(t, x(t)), \qquad t \in [a, b] \text{ a.e..}$$

The proof is complete. □

6.3 Necessary Conditions of Optimality

In this section we present, discuss and prove necessary conditions of optimality of the Pontryagin type for problem (P) stated in the first section of this chapter which also included the required assumptions on the data of the problem.

It is well known that necessary conditions are guaranteed to provide meaningful information if the solution to the problem exists and if the conditions do not degenerate. In this thesis, we are not concerned with either type of results since they would extend the volume of work of this thesis.

However, in (Baleanu *et al.* [28]) mild sufficient conditions for the existence of solution to fractional boundary value problem are given in Theorem 3.1 which corresponds to "fixing" the control function in the dynamics of the system which, in this chapter, are given in the form of a differential inclusion. As mentioned earlier, in this chapter we extend the maximum principle in Chapter 3 of Clarke's [43] to the fractional context having in mind the derivation of the main result of Chapter 7.

Theorem 6.3. *Let x be admissible fractional trajectory that solves the problem (P), and the assumptions (H1)-(H5) be satisfied. Then, there exists a multiplier $[p, \lambda, \xi, \mu, \gamma]$, such that $\lambda + \|\mu\| > 0$, where $\|\mu\| = \frac{1}{\Gamma(\alpha+1)} \int_a^b |\mu(dt)^\alpha|$ denotes the Radon measure norm, $\lambda > 0$ is a scalar, $p\colon [a,b] \to \mathbb{R}^n$ is a fractional adjoint function, ξ is a point in $\mathbb{R}^n$, $\mu(\cdot)$ is a non-negative Radon measure supported on the set*

$$S := \{t \in [a,b] \colon \partial_x^> h(t, x(t)) \neq \phi\},$$

and $\gamma\colon [a,b] \to \mathbb{R}^n$ is a measurable function. Then,

(i) $\xi \in \partial g(x(b))$;

(ii) *for almost all $t \in [a,b]$,*

$$\left(-p^{(\alpha)}(t), x^{(\alpha)}(t) \right) \in \partial_{x,p} H \left(t, x(t), p(t) + \frac{1}{\Gamma(\alpha+1)} \int_{[a,t)} \gamma(\tau)(d\mu(\tau))^\alpha \right),$$

where ∂H refers to the generalized gradient (see Appendix C) of the Hamiltonian with respect to (x,p), for a fixed t;

(iii) *for some $r > 0$, we have*

$$p(a) \in r \partial d_C(x(a)),$$

and

$$-p(b) - \frac{1}{\Gamma(\alpha+1)} \int_{[a,b]} \gamma(\tau)(d\mu(\tau))^\alpha \in \lambda \xi,$$

where $d_C(\cdot)$ is Euclidean distance function for C.

(iv) $\gamma(t) \in \partial_x^> h(t, x(t)) \quad \mu - a.e.,$

$\partial_x^> h(t, x(t))$ is a certain subset of the usual Clarke's generalized gradient of $h(t, x(t))$ with respect to x, for a fixed t (see Appendix C), defined by

$$\partial_x^> h(t, x(t)) = co\{\gamma = \lim_{i \to \infty} \gamma_i \colon \gamma_i \in \partial_x h(t_i, x_i), (t_i, x_i) \to (t, x), h(t_i, x_i) > 0\}.$$

Proof. The proof is organized in several steps. First, the original problem is modified by enlarging the velocity set in order to ensure the feasibility of perturbed trajectories. Second, we construct a sequence of auxiliary problems in which the various types of constraints are removed, and additional nonsmooth terms designed to properly penalize the violation of the removed constraints are added to the cost function. The optimization problems of the obtained sequence are simpler but require the application of Ekeland's variational principle. Third, we compute the generalized gradients for the perturbed problem and application of Fermat principle. Finally, we will express the results in terms of the original data in the Hamiltonian form.

Step 1. Let a multifunction $F(t,x)$ be defined in Ω, and for some arbitrary $\delta > 0$, consider the δ tube about x as defined as before and denoted by $T(x;\delta)$. Now, for any β, we define a new multifunction $F_\beta(t,\bar{x})$ in the closure tube $T(x;\delta/2)$ contained in Ω_δ such that $F_\beta(t,\bar{x}) = F(t,\bar{x}) + \beta B$. Therefore, when β goes to zero then the new multifunction $F_\beta(t,\bar{x})$ goes to the original multifunction $F(t,\bar{x})$. Furthermore, we define a fractional trajectory $y(\cdot)$ satisfying

$$\begin{aligned} y^{(\alpha)}(t) &\in F_\beta(t,y(t)), \\ y(a) &\in C, \\ (t,y(t)) &\in \Omega_\delta, \quad t \in [a,b]. \end{aligned}$$

Denote the set that contains all such fractional trajectories $y(t)$ by Λ_β. Then, for any small $\varepsilon > 0$ the function $\psi_\varepsilon(y)$ defined by

$$\psi_\varepsilon(y) = \max\left\{g(y(b)) - g(x(b)) + \varepsilon^2, \theta(y)\right\},$$

where the function $\theta(y)$ is given by

$$\theta(y) = \max_{a \le t \le b}\left\{h^+(t,y(t))\right\},$$

with $h^+(t,y(t)) = \max\{0, h(t,y(t))\}$. We consider a metric function $\Delta_\alpha(\cdot,\cdot)$ defined by

$$\Delta_\alpha(y,z) = \frac{1}{\Gamma(\alpha+1)}\int_a^b |y(t)-z(t)|\,(dt)^\alpha + |y(a)-z(a)|\,.$$

Step 2. Here, we apply penalization technique to incorporate the dynamic inclusion, and state constraint as penalty terms in the cost function. Since the solution to the penalized problem is not known, we need to apply a variational principle, and we choose the one due to Ivar Ekeland ([55]). Since for any positive $\beta < \varepsilon$, we have $\psi_\varepsilon(x) = \varepsilon^2$, it follows that

$$\psi_\varepsilon(x) \le \inf_{\Lambda_\beta} \psi_\varepsilon + \varepsilon^2.$$

Thus, Ekeland's Theorem asserts that there exists an element $z \in \Lambda_\beta$ that minimizes $\psi_\varepsilon(y) + \varepsilon\Delta_\alpha(y,z)$ over $y \in \Lambda_\beta$ such that

$$\Delta_\alpha(x,z) \le \varepsilon, \quad \psi_\varepsilon(z) \le \varepsilon^2.$$

To conclude this step, we need to show the following lemma.

Lemma 6.1. *For some $\delta > 0$, among all $y \in \Lambda_\beta$ satisfying $\|y - z\| < \delta$, there is a fractional trajectory z that minimizes*

$$\psi_\varepsilon(y) + \varepsilon\Delta_\alpha(y,z) + R_1 d_C(y(a)) + R_2 \boldsymbol{d}^\alpha(y, F_\beta),$$

where $\boldsymbol{d}^\alpha(\cdot,\cdot)$ is defined as before, $R_1 = (L_1 + \varepsilon L_2)$, $R_2 = (L_3 + \varepsilon L_4)$, and L_1, L_2, L_3, L_4 are given by

$$\begin{aligned}
L_1 &= \max\{K_g, K_h\}\left[K \ln_\alpha(K) + 1\right], \\
L_2 &= \frac{\left(\frac{(b-a)^\alpha}{\Gamma(\alpha+1)} + 1\right)}{\max\{K_g, K_h\}} L_1, \\
L_3 &= K \max\{K_g, K_h\}, \\
L_4 &= \left(\frac{(b-a)^\alpha}{\Gamma(\alpha+1)} + 1\right) K,
\end{aligned}$$

where K, K_g and K_h are defined before in the hypotheses (H4), (H1) and (H5), respectively.

Remark that this lemma states that there is an optimization problem without constraints that can be regarded as a perturbation of (P) and for which we have a solution $z \in \Lambda_\beta$.

Proof Lemma 6.1. Suppose that this lemma is false, then there is a sequence of fractional trajectories $\{y_j\}$ converging to z for which the expression in the lemma is less than its value at z which is $\psi_\varepsilon(z)$. Let $c_j \in C$ such that

$$d_C(y_j(a)) = |y_j(a) - c_j|, \tag{6.17}$$

and let $\bar{y}_j$ be the fractional trajectory defined by

$$\bar{y}_j(t) = y_j(t) + c_j - y_j(a). \tag{6.18}$$

From the Lipschitz condition for associated function $d_{F(t,\cdot)}(\cdot)$ (see Proposition 6.1). Then, we have

$$\left| d_{F_\beta(t,\bar{y}_j(t))}(\bar{y}_j^{(\alpha)}(t)) - d_{F_\beta(t,y_j(t))}(y_j^{(\alpha)}(t)) \right| \le k(t)\,|\bar{y}_j(t) - y_j(t)| + \left| \bar{y}_j^{(\alpha)}(t) - y_j^{(\alpha)}(t) \right|. \tag{6.19}$$

By using (6.18) and the fact that $\bar{y}^{(\alpha)}(t) = y^{(\alpha)}(t)$, we conclude that

$$d_{F_\beta(t,\bar{y}_j)}(\bar{y}_j^{(\alpha)}(t)) \leq d_{F_\beta(t,y_j)}(y_j^{(\alpha)}(t)) + k(t)\left|\bar{y}_j(t) - y_j(t)\right|.$$

By applying the fractional Jumarie integral operator for all terms of this inequality

$$\frac{1}{\Gamma(\alpha+1)}\int_a^b d_{F_\beta(t,\bar{y}_j(t))}(\bar{y}_j^{(\alpha)}(t))(dt)^\alpha \leq \frac{1}{\Gamma(\alpha+1)}\int_a^b d_{F_\beta(t,y_j(t))}(y_j^{(\alpha)}(t))(dt)^\alpha$$

$$+\frac{1}{\Gamma(\alpha+1)}\int_a^b k(t)\left|\bar{y}_j(t) - y_j(t)\right|(dt)^\alpha.$$

From (6.18), (6.17) and Definition 6.2, we have

$$\boldsymbol{d}^\alpha(\bar{y}_j, F_\beta) \leq \boldsymbol{d}^\alpha(y_j, F_\beta) + d_C(y_j(a))\frac{1}{\Gamma(\alpha+1)}\int_a^b k(t)(dt)^\alpha.$$

Let $K = E_\alpha(\frac{1}{\Gamma(\alpha+1)}\int_a^b k(t)(dt)^\alpha)$, then, $\ln_\alpha(K) = \frac{1}{\Gamma(\alpha+1)}\int_a^b k(t)(dt)^\alpha$, (here, $\ln_\alpha(\cdot)$ denotes the inverse function of the Mittag-Leffler function $E_\alpha(\cdot)$, for more information about the relations between Mittag-Leffler function $E_\alpha(\cdot)$ and the inverse $\ln_\alpha(\cdot)$, see Appendix A). Then, we have

$$\boldsymbol{d}^\alpha(\bar{y}_j, F_\beta) \leq \boldsymbol{d}^\alpha(y_j, F_\beta) + d_C(y_j(a))\ln_\alpha(K). \tag{6.20}$$

For j is sufficiently large, then, according to Theorem 6.1, there exists a fractional trajectory $z_j \in F_\beta$ such that $z_j(a) = \bar{y}_j(a) = c_j \in C$, and

$$\left\|z_j - \bar{y}_j\right\| \leq \frac{1}{\Gamma(\alpha+1)}\int_a^b \left|z_j^{(\alpha)} - \bar{y}_j^{(\alpha)}\right|(dt)^\alpha \leq K\boldsymbol{d}^\alpha(\bar{y}_j, F_\beta). \tag{6.21}$$

Furthermore, we can write

$$\left\|z_j - y_j\right\| \leq \left\|z_j - \bar{y}_j\right\| + \left\|\bar{y}_j - y_j\right\|.$$

Now, from (6.18) and (6.21), we have

$$\left\|z_j - y_j\right\| \leq K\boldsymbol{d}^\alpha(\bar{y}_j, F_\beta) + d_C(y_j(a)),$$

and by (6.20), we conclude that

$$\left\|z_j - y_j\right\| \leq \left[1 + K\ln_\alpha(K)\right]d_C(y_j(a)) + K\boldsymbol{d}^\alpha(y_j, F_\beta). \tag{6.22}$$

Since, from the triangle inequality and by adding and subtracting the quantities $\psi_\varepsilon(y_j)$, we obtain

$$\psi_\varepsilon(z_j) + \varepsilon\Delta_\alpha(z_j, z) \leq \psi_\varepsilon(y_j) + \varepsilon\Delta_\alpha(y_j, z) + \varepsilon\Delta_\alpha(y_j, z_j) + \psi_\varepsilon(z_j) - \psi_\varepsilon(y_j), \tag{6.23}$$

Then, by using the definition of $\Delta_\alpha(\cdot,\cdot)$, the characterization of the function $\psi_\varepsilon(\cdot)$, and Lipschitz property for the functions $g(\cdot)$, $\theta(.)$ we conclude that

$$\psi_\varepsilon(z_j) - \psi_\varepsilon(y_j) \leq \max\{K_g, K_h\}\|z_j - y_j\|,$$

$$\begin{aligned}
\varepsilon\Delta_\alpha(y_j, z_j) &= \frac{\varepsilon}{\Gamma(\alpha+1)}\int_a^b |y_i(t) - z_i(t)|\,(dt)^\alpha + \varepsilon\,|y_i(a) - z_i(a)| \\
&\leq \frac{\varepsilon(b-a)^\alpha}{\Gamma(\alpha+1)}\|z_j - y_j\| + \varepsilon\|z_j - y_j\| \\
&\leq \varepsilon\left(\frac{(b-a)^\alpha}{\Gamma(\alpha+1)} + 1\right)\|z_j - y_j\|.
\end{aligned}$$

By substituting in (6.23), we have

$$\psi_\varepsilon(z_j) + \varepsilon\Delta_\alpha(z_j, z) \leq \psi_\varepsilon(y_j) + \varepsilon\Delta_\alpha(y_j, z) + \left[\varepsilon\left(\frac{(b-a)^\alpha}{\Gamma(\alpha+1)} + 1\right) + \max\{K_g, K_h\}\right]\|z_j - y_j\|. \quad (6.24)$$

By using (6.22), we conclude that

$$\begin{aligned}
\psi_\varepsilon(z_j) + \varepsilon\Delta_\alpha(z_j, z) &\leq \psi_\varepsilon(y_j) + \varepsilon\Delta_\alpha(y_j, z) + R_1 d_C(y_j(a)) + R_2\boldsymbol{d}^\alpha(y_j, F_\beta) \\
&< \psi_\varepsilon(z),
\end{aligned}$$

where $R_1 = (L_1 + \varepsilon L_2)$, $R_2 = (L_3 + \varepsilon L_4)$, and L_1, L_2, L_3, L_4 are defined above. Then, this is a contradiction of that z is the optimal solution of $\psi_\varepsilon(\cdot) + \varepsilon\Delta_\alpha(\cdot, z)$.
Lemma 6.1 is proved. □

We now compute the generalized gradient for the functions in the Lemma 6.1.

Step 3. In this step, we calculate generalized gradients for the perturbed problem. The functions in Lemma 6.1 have $y = 0$ as a local minimizer of $\psi_\varepsilon(z+y) + \varepsilon\Delta_\alpha(z+y, z) + R_1 d_C(y(a) + z(a)) + R_2\boldsymbol{d}^\alpha(y + z, F_\beta)$. Thus, by Fermat's principle, we have that

$$0 \in \partial_y\{\psi_\varepsilon(z) + \varepsilon\Delta_\alpha(z, z) + R_1 d_C(z(a)) + R_2\boldsymbol{d}^\alpha(z, F_\beta)\},$$

and, from the sum generalized calculus rule, that

$$0 \in \partial_y\psi_\varepsilon(z) + \varepsilon\partial_y\Delta_\alpha(z, z) + R_1\partial_y d_C(z(a)) + R_2\partial_y\boldsymbol{d}^\alpha(z, F_\beta). \quad (6.25)$$

If $f(\cdot)$ is any Lipschitz function on $\mathbb{R}^n$, any element ξ of the generalized gradient of the mapping $y \to f(y(a))$ at y_0 is represented by an element $\xi_0 \in \partial f(y_0(a))$, so that $\xi(y) = \langle\xi_0, y(a)\rangle$ for all y (Clarke [43]). It follows that, if the state constraint is inactive, that is, $\theta(z) < \psi_\varepsilon(z)$, the function $\psi_\varepsilon(y)$ becomes $\psi_\varepsilon(y) = g(y(b)) - g(x(b)) + \varepsilon^2$, and the generalized gradient the map $y \to g(z(b)+y(b))$ at 0 is some element ξ_1 of $\partial g(z(b))$ and, thus, we have

$$\xi(y) = \langle \xi_1, y(b) \rangle . \tag{6.26}$$

If $\theta(z) > 0$, then ξ is an element of $\partial\theta(z)$ defined by

$$\xi(y) = \frac{1}{\Gamma(\alpha+1)} \int_{[a,b]} \langle \gamma(t), y(t) \rangle \, (d\mu(t))^\alpha, \tag{6.27}$$

where $\mu(\cdot)$ is a Radon measure on $[a,b]$ support on the points in time at which the constraint becomes active and $\gamma(t) \in \partial_x^{>} h(t, z(t))$, μ a.e.. By putting together all possibilities, we have that the generalized gradient ξ of the function $\partial_y \psi_\varepsilon(z)$ satisfies

$$\xi(y) = \lambda \langle \xi_1, y(b) \rangle + \frac{(1-\lambda)}{\Gamma(\alpha+1)} \int_{[a,b]} \langle \gamma(t), y(t) \rangle \, (d\mu(t))^\alpha. \tag{6.28}$$

Similarly, any element ξ of $R_1 \partial_y d_C(z(a))$ is represented by an element $\xi_0 \in R_1 \partial d_C(z(a))$, so that

$$\xi(y) = \langle \xi_0, y(a) \rangle . \tag{6.29}$$

Any element ξ of $R_2 \partial_y \boldsymbol{d}^\alpha(z, F_\beta)$ is represented by an element $(q,s) \in R_2 \partial \boldsymbol{d}^\alpha(z, F_\beta)$, as it follows from the generalized calculus rule (Clarke [43]), we have

$$\xi(y) = \frac{1}{\Gamma(\alpha+1)} \int_a^b \langle q(t), y(t) \rangle \, (dt)^\alpha + \frac{1}{\Gamma(\alpha+1)} \int_a^b \left\langle s(t), y^{(\alpha)}(t) \right\rangle (dt)^\alpha. \tag{6.30}$$

Similarly, any element ξ of $\varepsilon \partial_y \Delta_\alpha(\cdot,\cdot)$ corresponds to the function $r(t)$ with $r(t) \in \varepsilon B$ and a point $r_0 \in \varepsilon B$ such that

$$\xi(y) = \langle r_0, y(a) \rangle + \frac{1}{\Gamma(\alpha+1)} \int_a^b \langle r, y \rangle \, (dt)^\alpha. \tag{6.31}$$

Lemma 6.2. *From (6.25), there is a generalized subgradient satisfying*

$$\begin{aligned} &\langle \xi_0, y(a) \rangle + \lambda \langle \xi_1, y(b) \rangle + \langle r_0, y(a) \rangle \\ &\qquad + \tfrac{1}{\Gamma(\alpha+1)} \textstyle\int_a^b \langle q, y \rangle \, (dt)^\alpha + \tfrac{(1-\lambda)}{\Gamma(\alpha+1)} \textstyle\int_{[a,b]} \langle \gamma(t), y(t) \rangle \, (d\mu(t))^\alpha \\ &\qquad\qquad + \tfrac{1}{\Gamma(\alpha+1)} \textstyle\int_a^b \langle s, y^{(\alpha)} \rangle \, (dt)^\alpha + \tfrac{1}{\Gamma(\alpha+1)} \textstyle\int_a^b \langle r, y \rangle \, (dt)^\alpha = 0. \end{aligned}$$

Now, we are going to right this equality in terms of the data of the problem.

$$\langle \xi_0 + r_0, y(a)\rangle + \langle \lambda\xi_1, y(b)\rangle$$

$$+\tfrac{1}{\Gamma(\alpha+1)} \int_a^b \langle q + r + (1-\lambda)\gamma(t)\mu, y\rangle \, (dt)^\alpha$$

$$+\tfrac{1}{\Gamma(\alpha+1)} \int_a^b \left\langle s, y^{(\alpha)}\right\rangle (dt)^\alpha = 0.$$

By using the Jumarie fractional integration by parts (see Chapter 2), we have

$$\langle \xi_0 + r_0, y(a)\rangle + \langle \lambda\xi_1, y(b)\rangle + \tfrac{1}{\Gamma(\alpha+1)} \int_a^b \langle q + r + (1-\lambda)\gamma(t)\mu, y\rangle \, (dt)^\alpha$$

$$+\{y(b)s(b) - y(a)s(a)\} - \tfrac{1}{\Gamma(\alpha+1)} \int_a^b \left\langle s^{(\alpha)}, y\right\rangle (dt)^\alpha = 0.$$

Consequently,

$$\langle \xi_0 + r_0 - s(a), y(a)\rangle + \langle \lambda\xi_1 + s(b), y(b)\rangle$$

$$+\tfrac{1}{\Gamma(\alpha+1)} \int_a^b \left\langle q + r + (1-\lambda)\gamma(t)\mu - s^{(\alpha)}, y\right\rangle (dt)^\alpha = 0.$$

By applying the fractional Dubois-Reymond Lemma (see Appendix B), we conclude that

$$q + r + (1-\lambda)\gamma(t)\mu - s^{(\alpha)} = 0,$$

$$q + r + (1-\lambda)\gamma(t)\mu = s^{(\alpha)}.$$

By integrating, we obtain

$$\frac{1}{\Gamma(\alpha+1)} \int_a^t s^{(\alpha)} (d\tau)^\alpha = \frac{1}{\Gamma(\alpha+1)} \int_a^t (q+r)(d\tau)^\alpha + \frac{(1-\lambda)}{\Gamma(\alpha+1)} \int_{[a,t)} \gamma(t)\mu(d\tau)^\alpha,$$

and

$$s(t) = s(a) + \frac{1}{\Gamma(\alpha+1)} \int_a^t (q+r)(d\tau)^\alpha + \frac{(1-\lambda)}{\Gamma(\alpha+1)} \int_{[a,t)} \gamma(t)\mu(d\tau)^\alpha,$$

where,

$$\begin{aligned} s(a) &= \xi_0 + r_0, \\ -s(b) &= \lambda \xi_1. \end{aligned}$$

Therefore,

$$s(t) = \xi_0 + r_0 + \frac{1}{\Gamma(\alpha+1)} \int_a^t (q+r)(d\tau)^\alpha + \frac{(1-\lambda)}{\Gamma(\alpha+1)} \int_{[a,t)} \gamma(\tau)(d\mu(\tau))^\alpha.$$

Now, if we define a function $p(\cdot)$ as

$$p(t) = \xi_0 + r_0 + \frac{1}{\Gamma(\alpha+1)} \int_a^t (q+r)(d\tau)^\alpha,$$

then $p(t)$ satisfies

$$\begin{aligned} p(a) &= s(a) = \xi_0 + r_0, \\ p(b) &= s(b) - \frac{(1-\lambda)}{\Gamma(\alpha+1)} \int_{[a,b]} \gamma(\tau)(d\mu(\tau))^\alpha \\ &= -\lambda\xi_1 - \frac{(1-\lambda)}{\Gamma(\alpha+1)} \int_{[a,b]} \gamma(\tau)(d\mu(\tau))^\alpha. \end{aligned}$$

Consequently, we will conclude that

$$[p^{(\alpha)} - r, p + \frac{(1-\lambda)}{\Gamma(\alpha+1)} \int_{[a,t)} \gamma(\tau)(d\mu(\tau))^\alpha] \in R_2 \partial \boldsymbol{d}^\alpha(z, F_\beta), \tag{6.32}$$

$$p(a) \in R_1 \partial d_C(z(a)) + \varepsilon B, \tag{6.33}$$

$$-p(b) - \frac{(1-\lambda)}{\Gamma(\alpha+1)} \int_{[a,b]} \gamma(\tau)(d\mu(\tau))^\alpha = \lambda\xi_1. \tag{6.34}$$

Step 4. In this last step of the proof, we express the obtained conditions in terms of the Hamiltonian function. To this end, we simply apply the following lemma.

Lemma 6.3. *Let* $(q,p) \in \partial K \boldsymbol{d}^\alpha(y, F_\beta)$ *and* $\boldsymbol{d}^\alpha(y, F_\beta) = 0$. *Then, we have*

$$(-q, v) \in \partial H(t, y, p) + \beta B.$$

Its proof can be found in Clarke [43], Lemma 4, P. 127, and it yields the following condition

$$(-p^{(\alpha)}(t), z^{(\alpha)}(t)) \in \partial H \left(t, z(t), p(t) + \frac{(1-\lambda)}{\Gamma(\alpha+1)} \int_{[a,t)} \gamma(\tau)(d\mu(\tau))^\alpha \right) + 2\varepsilon B. \tag{6.35}$$

Here, we used the bounds $|r| \leq \varepsilon$, $\beta \leq \varepsilon$, and ε is any positive parameter (sufficiently small). Obviously that (6.33), (6.34) and (6.35) are the necessary conditions of optimality of the perturbed problem approximating the original one. Now, the final effort of the proof consists in showing that these equations converges to the necessary conditions of optimality of the original problem as the perturbation goes to zero. We know from the step 2 that $\Delta_\alpha(z, x) \leq \varepsilon$, so z converges to x as $\varepsilon \to 0$. We may select a sequence of the perturbations ε converging to zero such that λ converges to $\lambda_0 \in [0, 1]$.

By using Radon-Nikodym Theorem (Clarke [43]), it can be shown that for a further subsequence, the measures η defined by $d\eta = (1 - \lambda)\gamma d\mu$ converge weak* to a measure η_0 of the form $d\eta_0 = \gamma_0 d\mu_0$, where μ_0 is the weak* limit of $(1-\lambda)\gamma$, and γ_0 is a measurable selection of $\partial_x^{>} h(t, x(t))$. In consequence, $\lambda_0 + \|\mu_0\| > 0$.

Let we define a multifunction F and a measurable function y_ε as follows

$$\begin{aligned} F(t,x,p) &= \left\{(-v,u) : (u,v) \in \partial H\left(t,x,p + \frac{1}{\Gamma(\alpha+1)} \int_{[a,t)} \gamma_0 (d\mu_0)^\alpha\right)\right\}, \\ y_\varepsilon &= \frac{(1-\lambda_\varepsilon)}{\Gamma(\alpha+1)} \int_{[a,t)} \gamma(d\mu)^\alpha - \frac{1}{\Gamma(\alpha+1)} \int_{[a,t)} \gamma_0 (d\mu_0)^\alpha. \end{aligned}$$

Thus, for each ε, we have

$$(z_\varepsilon^{(\alpha)}, p_\varepsilon^{(\alpha)}) \in F(t, z_\varepsilon, p_\varepsilon + y_\varepsilon) + 2\varepsilon B,$$

and that y_ε converges to zero as $\varepsilon \to 0$. By using Theorem 6.2, to deduce the convergence of $(z_\varepsilon, p_\varepsilon)$ to (x, p).

Consequently, we obtain the necessary conditions of optimality for the original problem and thus the proof of the Theorem 6.3 is complete. $\square$

Chapter 7

Maximum Principle for Fractional Optimal Control Problems with State Constraints

7.1 Introduction

In this chapter, we state, discuss and prove a Pontryagin maximum principle for fractional optimal-control problems (FOCPs) with state constraints and under weak assumptions imposed on the data of the problem. We consider FOCP with the dynamics involving Jumarie fractional derivatives with respect to time.

The technique we used in the proof of this chapter is based in the construction of a fractional dynamical control system (FDCS) in such a way that, for each extremal of the FOCP there corresponds a process of the FDCS such that the endpoint of its fractional trajectory is on the boundary of the image of the attainable set by a certain Lipschitzian function of the state variable of the original system. The development of the FOCP must also be such that it enables the derivation of the necessary conditions for optimality in the form of a Pontryagin maximum principle for the FOCP from the characterization of boundary of the considered function of the attainable set of FDCS.

Definition 7.1. *Let the set C be a subset of $\mathbb{R}^n$, and $x(\cdot)$ is an admissible fractional trajectory associated to a control $u(\cdot)$. Assume that $[a, b]$ a given interval where $x(\cdot)$ satisfies the initial condition $x(a) \in C$. The set of all points $x(b)$ obtained by considering any admissible control function is called attainable set from C, and is denoted by $\mathcal{A}[C]$.*

The aims of this chapter is to study the FOCP in which the state constraints are present and give the necessary conditions for this problem.

7.2 The Problem Statement and Assumptions

In this section, we state the optimal-control problem for which the dynamics takes the form of a controlled Jumarie fractional differential equation with state constraints.

We consider the fractional optimal-control problem (P_C) as the following

$$
\begin{aligned}
(P_C)\,\text{Minimize} \quad & g(x(b)) \\
\text{subject to} \quad & x^{(\alpha)}(t) = f(t, x(t), u(t)), \quad t \in [a,b]\ \mathcal{L} - \text{a.e.}, && (7.1)\\
& x(a) \in C_0, \quad x(b) \in C_1, && (7.2)\\
& u \in \mathcal{U}, && (7.3)\\
& h(t, x(t)) \leq 0, \qquad \forall t \in [a,b], && (7.4)
\end{aligned}
$$

where $f : [a,b] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ is a fractional dynamics, $g : \mathbb{R}^n \to \mathbb{R}$ is a cost function, $h : [a,b] \times \mathbb{R}^n \to \mathbb{R}$ is a state constraint, $u : [a,b] \to \mathbb{R}^m$ is a measurable control satisfies the control constraints $u \in \mathcal{U}$, being $\mathcal{U}$ the set of measurable functions taking values on a set $\Omega(t)$ for each $t \in [a,b]$, and Ω is a set valued map taking on values on subsets of $\mathbb{R}^m$, C_0 and C_1 are a closed sets, $x(a)$ and $x(b)$ are initial and terminal point, respectively, and the operator (α) is Jumarie fractional derivative of order $\alpha \in (0,1]$ of the state variable with respect to time.

A maximum principle to be satisfied by the solutions to problem (P_C) are obtained under the following hypotheses:

(H1) The mapping $(t,u) \to f(t,x,u)$ is $\mathcal{L} \times \mathcal{B}-$measurable, where $\mathcal{L}$ and $\mathcal{B}$ denote the Lebesgue subset of $[a,b]$ and the Borel subset of $\mathbb{R}^n$, respectively. More $t \to f(t, x(t), u(t))$ is $\alpha-$integrable along any feasible control process.

(H2) For each $(t,u) \in \mathrm{Gr}(\Omega)$, there exists $\mathcal{L} \times \mathcal{B}-$measurable $k(t,u) : [a,b] \times \mathbb{R}^m \to \mathbb{R}$ defined in the $\mathrm{Gr}(\Omega)$ such that a function $f(t,\cdot,u)$ is Lipschitz of rank $k(t,u)$ and

$$|f(t,x_1,u) - f(t,x_2,u)| \leq k(t,u)\,|x_1 - x_2|\,.$$

(H3) $\mathrm{Gr}(\Omega)$ is $\mathcal{L} \times \mathcal{B}-$measurable, where $\mathrm{Gr}(\Omega)$ is a graph of the multifunction $\Omega : [a,b] \to \mathcal{P}(\mathbb{R}^m)$ defined by

$$\mathrm{Gr}(\Omega) := \{(t,u) \in [a,b] \times \mathbb{R}^m \colon u \in \Omega(t)\}\,.$$

(H4) The function $g(\cdot)$ is Lipschitz of rank K_g such that

$$|g(x) - g(\bar{x})| \leq K_g\,|x - \bar{x}|\,.$$

(H5) The function $h(\cdot,\cdot)$ is upper semicontinuous, and $h(t,\cdot)$ is Lipschitz of rank K_h for each $t \in [a,b]$ such that

$$|h(t,x) - h(t,\bar{x})| \leq K_h |x - \bar{x}| .$$

The Pontryagin function $H\colon [a,b] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ defined by

$$H(t,x,p,u) = \langle p, f(t,x,u) \rangle .$$

A pair (x,u) comprising an $\alpha-$absolutely continuous function x (the state fractional trajectory) and a measurable function u (the control), is called a feasible process of the problem (P_C), if satisfies all the constraints of the problem (P_C).

For FOCPs one may talk of local or global minimum, we say that x^* be a global minimum of (P_C), if it minimizes the objective function over all other feasible states $x \in \mathbb{R}^n$, and we say that x^* be a local minimum of (P_C), if it minimizes the objective function over all other feasible states $x \in \mathbb{R}^n$, in some neighborhood such that $|x - x^*| \leq \varepsilon$. Here, we restrict our discussion to minimizers in the context of Pontryagin type of minimum (Pontryagin *et al.* [127]). The conditions of the maximum principle selects only the control processes that are candidates to local maximum.

7.3 Maximum Principle of Optimality

Theorem 7.1. *Let the control process (x,u) be a solution to the problem (P_C), and assume that the assumptions (H1)–(H5) are satisfied. Then, there exists a fractional adjoint function $p\colon [a,b] \to \mathbb{R}^n$, a scalar $\lambda \geq 0$, a measurable function $\gamma(\cdot)$, a positive Radon measure $\mu(\cdot)$ supported on the set*

$$\{t \in [a,b] \colon h(t,x(t)) = 0\},$$

and a function $q(\cdot)$ defined by

$$q(t) = \begin{cases} p(t) + \frac{1}{\Gamma(\alpha+1)} \int_{[a,t)} \gamma(\tau)(d\mu(\tau))^{\alpha}, & t \in [a,b), \\ \\ p(t) + \frac{1}{\Gamma(\alpha+1)} \int_{[a,b]} \gamma(\tau)(d\mu(\tau))^{\alpha}, & t = b, \end{cases}$$

satisfying

(1) The adjoint equation

$$-p^{(\alpha)}(t) \in \partial_x H(t, x(t), q(t), u(t));$$

(2) The transversality condition

$$\begin{aligned} p(a) &\in N_{C_0}(x(a)), \\ -q(b) &\in \lambda \partial g(x(b)) + N_{C_1}(x(b)); \end{aligned}$$

(3) The control strategy $u\colon [a,b] \to \mathbb{R}^m$ *maximizes in* $\Omega(t)$ *the mapping*

$$v \to H\left(t, x(t), q(t), v\right);$$

(4) $\gamma(t) \in \partial_x^> h(t, x(t)) \quad \mu - a.e.$; *and*

(5) $|p| + \|\mu\| + \lambda > 0$.

Here, $\partial_x(\cdot)$ refers to the generalized gradient in the sense of Clarke with respect to x for fixed t, $N_{C_0}(\cdot)$ is the limiting normal cone in the sense of Mordukhovich [110], $\partial_x^> h(t, x(t))$ is a generalized gradient for the state constraint function defined in Theorem 6.3 Chapter 6, and $p^{(\alpha)}(\cdot)$ is Jumarie fractional derivative of the adjoint variable with respect to t of order $0 < \alpha \leq 1$.

The main idea of the proof is to construct an auxiliary dynamic control system in such a way that to each optimal-control process to (P_C) there corresponds a boundary control process to the auxiliary system. This ensures the non-triviality of the multiplier. Therefore, we are able to derive the necessary condition for optimal-control problem (P_C) from the characterization of a certain function of the attainable set of the auxiliary dynamical system. To achieve this, let $\tilde{s} = [s, s_1, s_0]$ denote points in $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}$,

$$\begin{aligned} \tilde{C} &= C_0 \times C_1 \times [0, \infty), \\ \tilde{f}(t, \tilde{s}(t), u(t)) &= \left(f(t, s(t), u(t)), 0, 0\right), \\ \theta(\tilde{s}) &= \left(g(s), s_1 - s\right), \\ \tilde{h}(t, \tilde{s}) &= h(t, s). \end{aligned}$$

Furthermore, suppose that $\tilde{x}(t) = [x, x_1, x_0]$ be admissible fractional trajectory satisfying

$$\begin{aligned} \tilde{x}^{(\alpha)}(t) &= \tilde{f}(t, \tilde{x}(t), u(t)), \\ \tilde{x}(a) &\in \tilde{C}, \\ \tilde{h}(t, \tilde{x}(t)) &\leq 0, \end{aligned}$$

where $\theta(\tilde{x}(b))$ is Lipschitz function lies in the boundary of $\theta(\mathcal{A}[\tilde{C}])$, in which $\mathcal{A}[\tilde{C}]$ is attainable set from $\tilde{C}$ at $t = b$ (see Definition 7.1). Now, we consider the following auxiliary results.

Lemma 7.1. *Let $\tilde{x}(t)$ be admissible fractional trajectory satisfies*

$$\begin{array}{rcl} \tilde{x}^{(\alpha)}(t) & = & \tilde{f}(t, \tilde{x}(t), u(t)), \\ \tilde{x}(a) & \in & \tilde{C}, \\ \tilde{h}(t, \tilde{x}(t)) & \leq & 0, \\ \theta(\tilde{x}(b)) & \in & bdy\ \theta(\mathcal{A}[\tilde{C}]), \end{array}$$

where bdy denotes the boundary. Then, there exists a vector ζ, a fractional adjoint function $\tilde{p}\colon [a,b] \to \mathbb{R}^n$, a measurable function $\gamma(\cdot)$, a positive Radon measure $\mu(\cdot)$ supported on the set

$$\{t \in [a,b] \colon \tilde{h}(t, \tilde{x}(t)) = 0\},$$

and $\tilde{q}(\cdot)$ defined by

$$\tilde{q}(t) = \begin{cases} \tilde{p}(t) + \frac{1}{\Gamma(\alpha+1)} \int_{[a,t)} \gamma(\tau)(d\mu(\tau))^{\alpha}, & t \in [a,b), \\ \\ \tilde{p}(t) + \frac{1}{\Gamma(\alpha+1)} \int_{[a,b]} \gamma(\tau)(d\mu(\tau))^{\alpha}, & t = b, \end{cases}$$

satisfying

(a) The adjoint equation

$$-\tilde{p}^{(\alpha)}(t) \in \partial_{\tilde{x}} H(t, \tilde{x}(t), \tilde{q}(t), u(t)),$$

(b) the transversality condition

$$\tilde{p}(a) \in N_{\tilde{C}}(\tilde{x}(a)),$$

$$\tilde{q}(b) \in \zeta \partial_{\tilde{x}} \theta(\tilde{x}(b)),$$

(c) maximum condition

$$\max_{v \in \Omega(t)} H\left(t, \tilde{x}(t), \tilde{q}(t), v\right) = H\left(t, \tilde{x}(t), \tilde{q}(t), u\right),$$

(d) $\gamma(t) \in \partial_{\tilde{x}}^{>} \tilde{h}(t, \tilde{x}(t)) \quad \mu - a.e.$,

(e) $\|\mu\| + |\zeta| > 0$.

Now, we show that the necessary conditions (1)-(5) for (P_C) are a consequence of the necessary conditions (a)-(e) of Lemma 7.1.

Let $\tilde{p} = (p, p_1, p_0)$, and, thus, $\tilde{q} = (q, q_1, q_0)$, where the relation between $\tilde{p}$ and $\tilde{q}$ are as mentioned earlier. Then, the Pontryagin function satisfies

$$H(t, \tilde{x}(t), \tilde{q}(t), u(t)) = \left\langle \tilde{q}(t), \tilde{f}(t, \tilde{x}(t), u(t)) \right\rangle.$$

From condition (a) of the Lemma 7.1, and the definitions of $\tilde{p}(t)$, $\tilde{q}(t)$, $\tilde{f}(t, \tilde{x}(t), u(t))$, and $H(t, \tilde{x}(t), \tilde{q}(t), u(t))$, we have

$$\left(-p^{(\alpha)}(t), -p_1^{(\alpha)}(t), -p_0^{(\alpha)}(t)\right) \in (q(t), q_1(t), q_0(t)) \begin{pmatrix} \partial_x f(t, x(t), u(t)) \\ 0 \\ 0 \end{pmatrix}.$$

Consequently, we obtain that

$$\begin{aligned} -p^{(\alpha)}(t) &\in q(t)\partial_x f(t, x(t), u(t)), \\ -p_1^{(\alpha)}(t) &= 0, \text{ and } -p_0^{(\alpha)}(t) = 0. \end{aligned}$$

By using the definition of Pontryagin function we have

$$-p^{(\alpha)}(t) \in \partial_x H(t, x(t), q(t), u(t)).$$

Then, condition (1) of Theorem 7.1 achieved. From the corollary C.1 of Appendix C and condition (b) of Lemma 7.1 we conclude that

$$p(a) \in N_{C_0}(x(a)), \quad p_1 \in N_{C_1}(x(b)), \quad \text{and } p_0 \leq 0.$$

Also, from condition (b) of the Lemma 7.1, we have that

$$\tilde{q}(b) \in \zeta \partial_{\tilde{x}} \theta(\tilde{x}(b)),$$

and, therefore, from the definitions of $\tilde{q}(\cdot)$, $\tilde{x}(\cdot)$, and $\theta(\cdot)$, by considering $\zeta = (\zeta_1, \zeta_2)$, we have

$$\begin{aligned} (q, q_1, q_0)(b) &\in (\zeta_1, \zeta_2)\partial_{(x, x_1, x_0)} \begin{pmatrix} g(x) \\ x_1 - x \end{pmatrix} (b) \\ &\in (\zeta_1, \zeta_2) \begin{pmatrix} \partial_x g(x(b)) & 0 & 0 \\ \{-I\} & \{I\} & 0 \end{pmatrix} \\ &\in (\zeta_1 \partial_x g(x(b)) - \zeta_2, \{\zeta_2\}, \{0\}), \end{aligned}$$

this means that

$$\begin{aligned} q(b) &\in \zeta_1 \partial_x g(x(b)) - \zeta_2, \\ q_1 &= \zeta_2. \end{aligned}$$

Then, by $\lambda = -\zeta_1$, we have condition (2) of the Theorem 7.1. We obtain the conditions (3), (4) and (5) by direct substitution of $\tilde{q}(\cdot)$, $\tilde{x}(\cdot)$ and ζ in the conditions (c), (d) and (e).

Since the necessary conditions of the Theorem 7.1 are a consequence of the necessary conditions in Lemma 7.1, now we only need to prove the Lemma 7.1. It is not difficult to see that the hypotheses of (P_C) in Theorem 7.1 migrate to the FDCS considered in Lemma 7.1. So, from now on, we assume that hypotheses (H1)-(H5) remain in force for the data of FDCS. The remaining of the proof consists in formulating an auxiliary family of optimal-control problems with dynamics given by fractional differential inclusions whose corresponding sequence of solutions converge to a boundary control process for the FDCS. This allows us to use the maximum principle proved in Chapter 6. In order to facilitate the proof two additional hypothesis will be temporarily considered:

Extra hypotheses

(IH1) For each $t \in [a, b]$, the set $\Omega(t)$ has a finite number of points.

(IH2) The function $\tilde{f}(t, \tilde{x}(t), v)$ is bounded by $\alpha-$integrable function $\sigma(\cdot)$ for all $v \in \Omega(t)$, $t \in [a, b]$ a.e., such that

$$\left|\tilde{f}(t, \tilde{x}(t), v)\right| \leq \sigma(t), \qquad k(t, v) \leq \sigma(t),$$

where $k(\cdot, \cdot)$ is the function whose mentioned before in the hypothesis (H2).

In the last step of the proof we will show that the results remain valid in the absence of these extra hypotheses.

Proof Lemma 7.1. The following steps are required to prove this lemma.

Step 1. Construction of an auxiliary family of fractional optimal-control problems associated with the original FDCS.

For this purpose, we start by constructing a set-up in which Ekeland's variational principle can be applied in order to characterize the sequence of auxiliary control processes approximating the boundary control process.

To achieve this goal, let $\varepsilon > 0$, such that the $\varepsilon-$tube $T(x; 2\varepsilon)$ is contained in $\bar{\Omega}$, where $\bar{\Omega} \subseteq [a, b] \times \mathbb{R}^n$, and V be the space of the feasible controls and initial condition, *i.e.*, the set of all pairs (v, z) where z is a point in $\tilde{C}$ and $v : [a, b] \to \mathbb{R}^m$ is a measurable control satisfying $v \in \mathcal{U}$ for which there is an admissible fractional trajectory $\tilde{y}(\cdot)$ satisfying the fractional differential equation

$$\tilde{y}^{(\alpha)}(t) = \tilde{f}(t, \tilde{y}(t), v(t)), \quad \tilde{y}(a) = z,$$

where (α) is Jumarie fractional derivative operator, $0 < \alpha \leq 1$, $t \in [a, b]$, and satisfies the state constraint

$$\tilde{h}(t, \tilde{y}(t)) \leq 0 \quad a.e..$$

Also, we need a complete metric space to achieve this step. For this, let any controls $v_1, v_2 \in V$ defined by

$$\delta(v_1, v_2) = \mathcal{L} - \text{meas}\{t \in [a, b] : v_1(t) \neq v_2(t)\},$$

for the points (v_1, z_1) and (v_2, z_2) in the set V, we provide the set V with the metric function Δ such that

$$\Delta((v_1, z_1), (v_2, z_2)) = \delta(v_1, v_2) + |z_1 - z_2|\,.$$

It is easy to verified that Δ is metric space on V, since $\Delta((v_1, z_1), (v_2, z_2)) \geq 0$, $\Delta((v_1, z_1), (v_2, z_2)) = \Delta((v_2, z_2), (v_1, z_1))$, and to show that

$$\Delta((v_1, z_1), (v_2, z_2)) \leq \Delta((v_1, z_1), (v_3, z_3)) + \Delta((v_3, z_3), (v_2, z_2)),$$

let $v_1, v_2, v_3 \in V$ such that

$$\{t\colon v_1 \neq v_2\} \subset \{t\colon v_1 \neq v_3\} \cup \{t\colon v_3 \neq v_2\},$$

$$\text{meas}\{t\colon v_1 \neq v_2\} \leq \text{meas}\{t\colon v_1 \neq v_3\} + \text{meas}\{t\colon v_3 \neq v_2\}.$$

Therefore,

$$\delta(v_1, v_2) \leq \delta(v_1, v_3) + \delta(v_3, v_2).$$

Now, we show that the space (V, Δ) is a complete metric space.

Lemma 7.2. *Let the sequence $\{(v_i, z_i)\} \in V$ be a Cauchy sequence. Then, there is an element $(v_0, z_0) \in V$ such that $\{(v_i, z_i)\}$ converges to an admissible pair (v_0, z_0).*

Proof. Since the sequence is Cauchy, it suffices to show that subsequence converges to (v_0, z_0) in V. It follows from [Ekeland [55], Lemma 7.2], we can extract a subsequence satisfying

$$\Delta\left(\{(v_i, z_i)\}, \{(v_{i+1}, z_{i+1})\}\right) \leq 2^{-i}.$$

Since

$$\left|\bigcup_{i \geq k}\{t\colon v_i(t) \neq v_{i+1}(t)\}\right| \leq 2^{1-k},$$

we have that $\{v_k(t)\}$ converges in measure to $v_0(t)$ such that

$$v_0(t) = v_k(t) \quad \forall t \notin \underset{i \geq k}{\cup} \{t \colon v_i(t) \neq v_{i+1}(t)\}.$$

Then, a control v_0 exists such that $\delta(v_i, v_0) \to 0$ for almost every t. Since $\mathbb{R}^n$ is complete, and $\tilde{C}$ is closed then, z_i converges to an element $z_0 \in \tilde{C}$ such that $|z_i - z_{i+1}| \leq 2^{-i}$. It remains to show that (v_0, z_0) lies in V. To do this, let $\tilde{y}_i(\cdot)$ be the fractional trajectory associated with $\{(v_i, z_i)\}$, $\Gamma(t, \tilde{y})$ be a set valued map defined by

$$\Gamma(t, \tilde{y}) = \{\tilde{f}(t, \tilde{y}, v_0(t))\}.$$

From the above, it is clear that the set $\mathcal{A}_i$, defined by

$$\mathcal{A}_i = \{t \in [a, b] : v_i(t) = v_0(t)\},$$

is such that $\mathcal{L} - \text{meas}(\mathcal{A}_i) \to (b - a)$, since

$$\tilde{y}_i^{(\alpha)}(t) = \{\tilde{f}(t, \tilde{y}_i(t), v_i(t))\}, \quad \forall t \in [a, b].$$

Therefore, we have

$$\tilde{y}_i^{(\alpha)}(t) \in \Gamma(t, \tilde{y}_i(t)), \quad \forall t \in \mathcal{A}_i.$$

Then, by applying Theorem 6.2, we conclude that there exists an $\alpha-$absolutely continuous function $\tilde{y}_0(t)$ such that $\tilde{y}_0(0) = z_0$, and $\tilde{y}_0^{(\alpha)}(t) = \tilde{f}(t, \tilde{y}_0(t), v_0(t))$. Thus, $\tilde{y}_0(t)$ is an admissible trajectory corresponding to (v_0, z_0).

Lemma 7.2 is proved. □

Now, we prove that there is a fractional trajectory $\tilde{y}_i(t)$ associated with a sequence $\{(v_i, z_i)\}$ converges to $\tilde{y}_0(t)$ associated with (v_0, z_0).

Lemma 7.3. *If $\{(v_i, z_i)\} \in V$ converges to $(v_0, z_0) \in V$, then $|\tilde{y}_i(t) - \tilde{y}_0(t)|$ converges to zero.*

Proof. We have $\tilde{y}_0^{(\alpha)}(t) - \tilde{f}(t, \tilde{y}_0(t), v_0(t)) = 0$ on the set $\mathcal{A}_i$. Since $\left|\tilde{y}_0^{(\alpha)}(t) - \tilde{f}(t, \tilde{y}_0(t), v_0(t))\right|$ is bounded, and $\mathcal{L} - \text{meas}(\mathcal{A}_i) \to (b - a)$. Moreover, we know from the results of Chapter 6 that $d_{\Gamma(t,\cdot)}(\cdot)$ is associated function to a multifunction $\Gamma(t, \cdot)$, and $d_{\Gamma(t,\tilde{y}_i)}(\tilde{y}_i^{(\alpha)}) = 0$, when $t \in \mathcal{A}_i$, it follows from the mentioned results of the Theorem 6.1 in Chapter 6, for any positive δ_i, for all i sufficiently large, $\boldsymbol{d}^{\alpha}(\tilde{y}_i, \Gamma) \leq \delta_i$. Then, there exists a fractional trajectory $y_i(t)$ for a multifunction $\Gamma(t, y_i(t))$ satisfying $y_i(a) = \tilde{y}_i(a) = z_i$, $|y_i(t) - \tilde{y}_i(t)| \leq K\delta_i$ for each t where K defined in Theorem 6.1 in Chapter 6, and $y_i^{(\alpha)}(t) = \tilde{f}(t, y_i(t), v_0(t))$ then, we have

$$\left|\tilde{y}_0^{(\alpha)}(t) - y_i^{(\alpha)}(t)\right| \quad = \quad \left|\tilde{f}(t, \tilde{y}_0(t), v_0(t)) - \tilde{f}(t, y_i(t), v_0(t))\right|$$

$$\leq \quad \sigma(t)\left|\tilde{y}_0(t) - y_i(t)\right|,$$

where $\sigma(t)$ is defined before in the extra hypotheses, and also we have

$$\left|\tilde{y}_0(a) - y_i(a)\right| = \left|z_0 - z_i\right|.$$

By applying fractional Gronwall inequality (see Appendix B), we conclude that

$$\left|\tilde{y}_0(t) - y_i(t)\right| \leq \left|z_0 - z_i\right| E_\alpha\left(\sigma(t)\Gamma(\alpha)(t-a)^\alpha\right),$$

where $E_\alpha(\cdot)$ is generalization Mittag-Leffler function (see Chapter 2). Therefore, we conclude that

$$\begin{aligned} \left|\tilde{y}_0(t) - \tilde{y}_i(t)\right| &\leq \left|\tilde{y}_0(t) - y_i(t)\right| + \left|y_i(t) - \tilde{y}_i(t)\right| \\ &\leq \left|z_0 - z_i\right| E_\alpha(\sigma(t)\Gamma(\alpha)(t-a)^\alpha) + K\delta_i. \end{aligned}$$

Then, for i large, we have $|\tilde{y}_i(t) - \tilde{y}_0(t)| \to 0$.

Lemma 7.3 is proved. $\square$

Now, for a positive integer i we choose a point ξ in $\theta(\tilde{x}(b)) + i^{-2}B$ where B is open ball centered at zero, such that $\xi \notin \theta(\mathcal{A}[\tilde{C}])$, let

$$G(\tilde{x}(b)) = \left|\xi - \theta((\tilde{x}(b)))\right|,$$

where $\theta(\cdot)$ lie in the boundary of $\theta(\mathcal{A}[\tilde{C}])$. Since the function $\tilde{y}(t) \to \tilde{f}(t, \tilde{y}(t), v(t))$ is Lipschitz with constant $\sigma(t)$, and $\tilde{y}(b)$ is a fractional trajectory associated with (v, z) lie in the boundary of the attainable set $\mathcal{A}[\tilde{C}]$. Then,

$$G(\tilde{y}(b)) = \left|\xi - \theta(\tilde{y}(b))\right|.$$

By applying Ekeland Theorem (see Appendix B), then for $G(\tilde{y}(b))$ is non-negative and $(u(t), \tilde{x}(a)) \in V$ satisfies

$$G(\tilde{x}(b)) \leq \inf_V G(\tilde{y}(b)) + i^{-2},$$

where $\tilde{x}(b)$ is a fractional trajectory associated with $(u(t), \tilde{x}(a))$ lie in the boundary of the attainable set $\mathcal{A}[\tilde{C}]$, for some $i^{-2} > 0$, there exists a point $(\hat{v}, \hat{z}) \in V$ such that

$$\Delta((u, \tilde{x}(a)), (\hat{v}, \hat{z})) \leq \frac{1}{i}, \tag{7.5}$$

and

$$G(\tilde{y}(b)) \leq G(\tilde{x}(b)),$$

where $\hat{y}(b)$ is a fractional trajectory corresponding to $(\hat{v}, \hat{z})$ lie in the boundary of the attainable set $\mathcal{A}[\tilde{C}]$, for all $(\hat{v}, \hat{z}) \neq (v, z)$ in the set V then,

$$G(\tilde{y}(b)) + i^{-1}\Delta((v, z), (\hat{v}, \hat{z})) \geq G(\hat{y}(b)).$$

So, we can simplify the results of this step in the following Lemma.

Lemma 7.4. *Let $(\tilde{y}(t), v(t))$ be admissible process on the interval $[a, b]$ satisfying*

$$\begin{aligned} \tilde{y}(a) &\in \tilde{C}, \\ \tilde{h}(t, \tilde{y}(t)) &\leq 0, \\ |\tilde{y}(t) - \tilde{x}(t)| &\leq \varepsilon. \end{aligned}$$

Then,

$$G(\tilde{y}(b)) + i^{-1}\Delta((v, z), (\hat{v}, \hat{z})) \geq G(\hat{y}(b)),$$

for all $(\hat{v}, \hat{z}) \neq (v, z) \in V$. Therefore, we have

$$|\xi - \theta(\tilde{y}(b))| + i^{-1}\delta(v, \hat{v}) + i^{-1} |\tilde{y}(a) - \hat{z}| \geq |\xi - \theta(\hat{y}(b))| .$$

Now, let us use the results obtained in Section 2 of Chapter 6 in the current context.

Step 2. In this step, we construct a fractional differential inclusion in order to take advantage of the maximum principle proved in the previous section. To do that, let the fractional states $Y(t)$ having three components as follows $Y(t) = [\tilde{y}_1(t), \tilde{y}_2(t), \tilde{y}_3(t)]$, and the multifunction $F(t, Y(t))$ defined by

$$F(t, Y(t)) := \{[\frac{v}{1 + |v|}, \chi_t(v), \tilde{f}(t, \tilde{y}_3, v)] : v \in V\},$$

where the first component of the multifunction $F(t, Y(t))$ is a term responsible by the compactification, *i.e.*, it ensures the convergence of limiting sequences even when the control constraint set can become unbounded, the second component is a term to penalize the deviation of the control with respect to the optimizing one as $\chi_t(v)$ is indicator function defined by

$$\chi_t(v) = \begin{cases} 1, & \text{if } v \neq \hat{v}, \\ 0, & \text{otherwise}, \end{cases}$$

and the last component of the multifunction is the usual dynamics. Suppose that the set C defined by

$$C := \{[\tilde{y}_1, \tilde{y}_2, \tilde{y}_3] : \tilde{y}_3 \in \tilde{C}\}.$$

Thus, any fractional trajectory $Y(t) = [\tilde{y}_1(t), \tilde{y}_2(t), \tilde{y}_3(t)]$ for a multifunction $F(t, Y(t))$ from the set C defined by

$$Y = [\beta_1 + \frac{1}{\Gamma(\alpha+1)} \int_a^t \frac{v}{1+|v|} (d\tau)^\alpha, \beta_2 + \frac{1}{\Gamma(\alpha+1)} \int_a^t \chi_\tau(v)(d\tau)^\alpha, \tilde{y}(t)], \tag{7.6}$$

where $(\tilde{y}(t), v(t))$ is an admissible control process such that $\tilde{y}(0) \in \tilde{C}$, β_1, β_2 are constant, and $\int_a^t (\cdot)(d\tau)^\alpha$ is a fractional Jumarie integral operator with $\alpha \in (0, 1]$. Now, we define two functions $M_1(\cdot)$, $M_2(\cdot)$ such that the sum of these functions is equivalent to the cost functional that is minimized by the boundary control process. Let $M_1(Y)$ and $M_2(Y)$ defined by

$$\begin{aligned} M_1(Y) &= i^{-1} \left| \tilde{y}_3 - \tilde{z} \right| - i^{-1} \tilde{y}_2, \\ M_2(Y) &= \left| \xi - \theta(\tilde{y}_3) \right| + i^{-1} \tilde{y}_2. \end{aligned}$$

Therefore, from the Lemma 7.4, the fractional trajectory $\hat{Y}(t) = [\hat{y}_1(t), \hat{y}_2(t), \hat{y}_3(t)]$ given by (7.6) with $v = \hat{v}$, $\tilde{y} = \hat{y}$, $\beta_1 = 0$, $\beta_2 = 0$ minimizes

$$M_2(Y(b)) + M_1(Y(a)),$$

over the fractional trajectories for the multifunction $F(t, Y(t))$ in the set C, and satisfy

$$\left| \tilde{y}_3(t) - \tilde{x}(t) \right| \leq \varepsilon, \tag{7.7}$$

$$\tilde{h}(t, \tilde{y}_3(t)) \leq 0. \tag{7.8}$$

Let,

$$\begin{aligned} M(Y(a), Y(b)) &= M_2(Y(b)) + M_1(Y(a)) \\ &= \left| \xi - \theta(\tilde{y}_3(b)) \right| + i^{-1} \left| \tilde{y}_3 - \tilde{z} \right|. \end{aligned}$$

Obviously, we have that

$$M(\hat{Y}(a), \hat{Y}(b)) = \left| \xi - \theta(\hat{y}(b)) \right|.$$

Let the function $\Phi(Y)$ defined by

$$\Phi(Y) = \max\{\tilde{h}^+(t, \tilde{y}_3(t))\},$$

where $\tilde{h}^+ = \max\{\tilde{h}, 0\}$.

Then, we conclude that the fractional trajectory $\hat{Y}(t)$ minimizes

$$\max\left\{M(Y(a),Y(b))-M(\hat{Y}(a),\hat{Y}(b)),\Phi(Y)\right\}, \tag{7.9}$$

over the fractional trajectories $Y(t)$ for the multifunction $F(t,Y(t))$ such that

$$Y(a)\in C, \tag{7.10}$$

with the constraints in (7.7), and (7.8) in force. When i is sufficiently large, (7.5) implies that any $\tilde{y}(t)$ near $\hat{y}(t)$ will automatically satisfy (7.7). So, $\hat{Y}(t)$ provides a strong local minimum (see Appendix D) for the function in (7.9) subject to the constraint (7.10).

Step 3. Now, we define the Hamiltonian $H(\cdot,\cdot,\cdot)$. For this purpose, let the adjoint fractional function $P(\cdot)$ that also has three components $[\tilde{p}_1(\cdot),\tilde{p}_2(\cdot),\tilde{p}_3(\cdot)]$, therefore, $Q(\cdot)=[\tilde{q}_1(\cdot),\tilde{q}_2(\cdot),\tilde{q}_3(\cdot)]$ where the relation between $\tilde{q}(\cdot)$ and $\tilde{p}(\cdot)$ as defined before. Then, the Hamiltonian defined by

$$H(t,Y,Q):=\max_{W\in F(t,Y)}\langle Q,W\rangle,$$

where Y, $F(t,Y)$ defined before. Therefore, $H(\cdot,\cdot,\cdot)$ is as follows

$$H(t,Y,Q):=\max_{v\in\mathcal{U}(t)}\left\{\frac{\langle\tilde{q}_1,v\rangle}{1+|v|}+\chi_t(v)\tilde{q}_2+\left\langle\tilde{q}_3,\tilde{f}(t,\tilde{y}_3,v)\right\rangle\right\}.$$

By applying the results obtained in Section 2 of Chapter 6, and by using the generalized gradient (see Appendix C) for the Hamiltonian $H(t,Y,Q)$, we have

$[-\tilde{p}_3^{(\alpha)}(t),0]\in co\{[D\tilde{q}_3(t),r]:D,\tilde{q}_3(t),$ and r are all the elements that can be defined as follows:

$$\tilde{q}_3(t)=\tilde{p}_3(t)+\frac{1}{\Gamma(\alpha+1)}\int_{[a,t)}\gamma(\tau)(d\mu(\tau))^\alpha,\quad D=\lim_{i\to\infty}\partial_{\tilde{y}}\tilde{f}(t,\tilde{y}_i,v_i),\quad r=\lim_{i\to\infty}\chi_t(v_i)\}.$$

From the definition of the indicator function, there are several values of r it that will be zero. Thus, for large i, we have only relevant sequences $\{v_i\}$ that satisfy $\{v_i\}=\hat{v}$, and as a consequence, $\tilde{y}_i(t)$ converges to $\tilde{y}(t)$. Then, by setting $\tilde{p}_3(t)=\tilde{p}(t)$, and $\tilde{q}_3(t)=\tilde{q}(t)$ we conclude that

$$-\tilde{p}^{(\alpha)}(t)\in\tilde{q}(t)\partial_{\tilde{y}}\tilde{f}(t,\hat{y}(t),\hat{v}(t)),$$

by using the definition of Pontryagin function we have

$$-\tilde{p}^{(\alpha)}(t)\in\partial_{\tilde{y}}H(t,\hat{y}(t),\tilde{q}(t),\hat{v}(t)). \tag{7.11}$$

Note that in Theorem 6.3 of Chapter 6 the transversality conditions and other component of the differential inclusion imply that $\tilde{p}_1 = 0$, and $\tilde{p}_2$ is constant, and the same happens to $\tilde{q}_1$, and $\tilde{q}_2$, respectively.

Likewise, as the Theorem 6.3 in Chapter 6 there is a non-negative $\lambda > 0$ such as

$$\tilde{p}_3(a) \in r\partial_{\tilde{y}_3} d_C(\hat{Y}(a)) + \lambda\partial_{\tilde{y}_3} M_1(\hat{Y}(a)),$$

$$\tilde{p}_3(a) \in r\partial_{\tilde{y}_3} d_{\tilde{C}}(\hat{y}(a)) + i^{-1}B.$$

As $i \to \infty$, then from (7.5) and the definition of the metric $\Delta(\cdot,\cdot)$, we have that the measure of the set $\{t\colon \hat{v}(t) \neq u(t)\}$ goes to zero and $|\hat{y}(a) - \tilde{x}(a)| \to 0$. It follows from Lemma 7.3, that $\hat{y}(\cdot)$ converges to $\tilde{x}(\cdot)$. Then, the last equation (with $\hat{y}(\cdot) = \tilde{x}(\cdot)$, $\hat{v}(\cdot) = u(\cdot)$, and $\tilde{p}_3(\cdot) = \tilde{p}(\cdot)$) becomes

$$\tilde{p}(a) \in r\partial d_{\tilde{C}}(\tilde{x}(a)).$$

From Proposition C.2 in Appendix C, we have

$$\tilde{p}(a) \in N_{\tilde{C}}(\tilde{x}(a)). \tag{7.12}$$

Similarly,

$$-\tilde{p}_3(b) - \frac{1}{\Gamma(\alpha+1)}\int_{[a,b]} \gamma(\tau)(d\mu(\tau))^{\alpha} \in \lambda\partial_{\tilde{y}_3} M_2(\hat{Y}(b)),$$

since $\tilde{q}_3(b) = \tilde{p}_3(b) + \frac{1}{\Gamma(\alpha+1)}\int_{[a,b]} \gamma(\tau)(d\mu(\tau))^{\alpha}$, therefore,

$$-\tilde{q}_3(b) \in \lambda\partial_{\tilde{y}_3} M_2(\hat{Y}(b)),$$

$$-\tilde{q}_3(b) \in \lambda\partial_{\tilde{y}_3} G(\hat{y}(b)),$$

where $G(\cdot)$ is defined before. We know from above $\xi \notin \theta(\mathcal{A}[\tilde{C}])$, therefore, $\xi \neq \theta(\hat{y}(b))$, then the distance $G(\hat{y}(b)) \neq 0$. So, by using Theorem C.2 in Appendix C, we have

$$-\tilde{q}_3(b) \in \lambda\partial_{\tilde{y}_3} G(\hat{y}(b))\partial_{\tilde{y}_3}\theta(\hat{y}(b)).$$

From Proposition C.3 in Appendix C, we have

$$\zeta = \frac{-\lambda(\theta(\hat{y}(b)) - \xi)}{|\theta(\hat{y}(b)) - \xi|}.$$

Since $\|\mu\| + \lambda > 0$, then we have

$$\|\mu\| + |\zeta| > 0.$$

Therefore, we conclude that

$$\tilde{q}(b) \in \zeta\partial_{\tilde{y}}\theta(\hat{y}(b)).$$

By using the same notation when $i \to \infty$ which stated before for $\tilde{p}(a)$, we have

$$\tilde{q}(b) \in \zeta \partial_{\tilde{x}} \theta(\tilde{x}(b)). \tag{7.13}$$

Finally, since the components of the Hamiltonian inclusion satisfy $\tilde{q}_1(\cdot) = 0$, and $\tilde{q}_2(\cdot)$ is constant, then the maximization defined by

$$\left\langle \tilde{q}_3(t), \tilde{f}(t, \hat{y}(t), \hat{v}) \right\rangle \geq \max_{v \in \Omega(t)} \left\langle \tilde{q}_3(t), \tilde{f}(t, \hat{y}(t), v) \right\rangle.$$

By using the notation when $i \to \infty$, and $\hat{y}(\cdot) = \tilde{x}(\cdot)$, $\hat{v}(\cdot) = u(\cdot)$, and $\tilde{q}_3(\cdot) = \tilde{q}(\cdot)$, we have

$$\left\langle \tilde{q}(t), \tilde{f}(t, \tilde{x}(t), u) \right\rangle = \max_{v \in \Omega(t)} \left\langle \tilde{q}(t), \tilde{f}(t, \tilde{x}(t), v) \right\rangle.$$

So, by applying the Pontryagin form we conclude that

$$H(t, \tilde{x}(t), \tilde{q}(t), u) = \max_{v \in \Omega(t)} H(t, \tilde{x}(t), \tilde{q}(t), v). \tag{7.14}$$

Step 4. Finally, in order to complete the proof we will remove the extra hypotheses (IH1), (IH2) by showing that the results obtained above are valid without them. Let the hypothesis (IH2) be in force but the hypothesis (IH1) be absent. For this, let

$$\eta(t) = \tilde{p}(t) E_\alpha({}_aJ_t^\alpha \tilde{f}_{\tilde{x}}), \tag{7.15}$$

where ${}_aJ_t^\alpha(\cdot)$ is Jumarie fractional integral operator (see Chapter 2), and $E_\alpha(\cdot)$ is Mittag-Leffler function (see Chapter 2). By differentiating both sides by fractional Jumarie derivative, we have

$$\eta^{(\alpha)}(t) = \tilde{p}^{(\alpha)}(t) E_\alpha({}_aJ_t^\alpha \tilde{f}_{\tilde{x}}) + \tilde{p}(t) \tilde{f}_{\tilde{x}} E_\alpha({}_aJ_t^\alpha \tilde{f}_{\tilde{x}}).$$

From the adjoint equation of the Lemma 7.1 (condition (a)) we have

$$-\tilde{p}^{(\alpha)} = (\tilde{p} + {}_aJ_{t;\mu}^\alpha \tilde{h}_{\tilde{x}}) \tilde{f}_{\tilde{x}},$$

therefore,

$$\begin{aligned} \eta^{(\alpha)}(t) &= -\tilde{p}(t) \tilde{f}_{\tilde{x}} E_\alpha({}_aJ_t^\alpha \tilde{f}_{\tilde{x}}) - ({}_aJ_{t;\mu}^\alpha \tilde{h}_{\tilde{x}}) \tilde{f}_{\tilde{x}} E_\alpha({}_aJ_t^\alpha \tilde{f}_{\tilde{x}}) + \tilde{p}(t) \tilde{f}_{\tilde{x}} E_\alpha({}_aJ_t^\alpha \tilde{f}_{\tilde{x}}) \\ &= -({}_aJ_{t;\mu}^\alpha \tilde{h}_{\tilde{x}}) \tilde{f}_{\tilde{x}} E_\alpha({}_aJ_t^\alpha \tilde{f}_{\tilde{x}}). \end{aligned}$$

By integrating both sides, we obtain

$$\tilde{p}(t) = \eta(a) E_\alpha(-{}_aJ_t^\alpha \tilde{f}_{\tilde{x}}) - {}_aJ_t^\alpha \left(({}_aJ_{t;\mu}^\alpha \tilde{h}_{\tilde{x}}) \tilde{f}_{\tilde{x}} \right),$$

from the Lipschitz condition for the function $\tilde{f}$, $\tilde{h}$ we have

$$\begin{aligned}
\left|\tilde{f}_{\tilde{x}}\right| &\leq k(t,u), \\
\left|\tilde{h}_{\tilde{x}}\right| &\leq K_{\tilde{h}}, \quad \text{and} \\
\|\mu\| &= \frac{1}{\Gamma(\alpha+1)} \int_a^t |\mu(d\tau^{\alpha})| \leq 1.
\end{aligned}$$

Here, $K_{\tilde{h}}$ is Lipschitz constant for the function $\tilde{h}(t,\cdot)$. Additionally, $|\tilde{p}(a)| = |\eta(a)| \leq 1$. Then, we conclude that

$$\begin{aligned}
|\tilde{p}(t)| &\leq |\eta(a)|\, E_{\alpha}({}_aJ_t^{\alpha} \left|\tilde{f}_{\tilde{x}}\right|) + {}_aJ_t^{\alpha}\left(({}_aJ_{t;|\mu|}^{\alpha} \left|\tilde{h}_{\tilde{x}}\right|) \left|\tilde{f}_{\tilde{x}}\right|\right), \\
&\leq E_{\alpha}({}_aJ_t^{\alpha} k(t,u)) + K_{\tilde{h}}\, {}_aJ_t^{\alpha} k(t,u).
\end{aligned}$$

The right-hand side from this inequality defined by

$$\tilde{M} := E_{\alpha}\left(\frac{1}{\Gamma(\alpha+1)} \int_a^b k(t,u)(dt)^{\alpha}\right) + K_{\tilde{h}}\left(\frac{1}{\Gamma(\alpha+1)} \int_a^b k(t,u)(dt)^{\alpha}\right).$$

Consequently,

$$|\tilde{p}(t)| \leq \tilde{M}.$$

Let $\{S_j\}$ be an increasing family of finite subset of $\tilde{M}B$ such that $\tilde{M}B \subset S_j + j^{-1}B$ for all j. For each $s \in S_j$, we select a measurable function $v_s(\cdot)$ such that $v_s(t) \in \Omega(t)$ a.e., and

$$H(t,\tilde{x}(t),s) \leq \left\langle s, \tilde{f}(t,\tilde{x}(t),v_s)\right\rangle + j^{-1}, \tag{7.16}$$

where

$$H(t,\cdot,s) := \sup\{\left\langle s, \tilde{f}(t,\cdot,v)\right\rangle : v \in \Omega(t)\}. \tag{7.17}$$

Let

$$\Omega_j(t) = \{v_s(t) \colon s \in S_j\} \cup \{u(t)\}.$$

Now, consider a new problem in which the multifunction $\Omega(t)$ is replaced by $\Omega_j(t)$. By assuming that the hypotheses are satisfied by the data for the new problem, our result yields multipliers $\tilde{p}$, μ, γ, and ζ with properties listed in Lemma 7.1, except that, now, the maximization of the Hamiltonian condition takes the form

$$\left\langle \tilde{q}, \tilde{x}^{(\alpha)}\right\rangle = H_j(t,\tilde{x},\tilde{q}), \tag{7.18}$$

where H_j defined by

$$H_j(t,x,s) = \max\{\left\langle s, \tilde{f}(t,x,v)\right\rangle : v \in \Omega_j(t)\}.$$

For each t at which condition (c) of the Lemma 7.1 is valid, choose $s \in S_j$ which that

$$\tilde{q} = \tilde{p} + \frac{1}{\Gamma(\alpha+1)} \int_{[a,t)} \gamma(\tau)\mu(d\tau)^{\alpha} \in s + j^{-1}B. \tag{7.19}$$

Since we continue to assume condition (IH2) of extra hypotheses, $\sigma(t)$ is Lipschitz constant for $H(t,\tilde{x}(t),\cdot)$ and $H_j(t,\tilde{x}(t),\cdot)$ such that

$$\left|H_j(t,\tilde{x}(t),s) - H_j(t,\tilde{x}(t),\tilde{q})\right| = \sigma(t)\left|s-\tilde{q}\right|.$$

From (7.19), we have

$$\begin{aligned} \left|H_j(t,\tilde{x}(t),s)\right| - \left|H_j(t,\tilde{x}(t),\tilde{q})\right| &\leq \sigma(t)\left|j^{-1}\right|, \\ \left|H_j(t,\tilde{x}(t),s)\right| - \sigma(t)\left|j^{-1}\right| &\leq \left|H_j(t,\tilde{x}(t),\tilde{q})\right|. \end{aligned}$$

By using the definition of $H_j(t,\tilde{x}(t),\cdot)$, we have

$$\left\langle s, \tilde{f}(t,\tilde{x}(t),v_s)\right\rangle - \sigma(t)\left|j^{-1}\right| \leq \left|H_j(t,\tilde{x}(t),\tilde{q})\right|.$$

By using (7.16), we have

$$\left|H_j(t,\tilde{x}(t),\tilde{q})\right| \geq \left|H(t,\tilde{x}(t),s)\right| - j^{-1} - \sigma(t)\left|j^{-1}\right|,$$

and by applying Lipschitz condition for $H(t,\tilde{x}(t),\cdot)$ we obtain

$$\left|H_j(t,\tilde{x}(t),\tilde{q})\right| \geq \left|H(t,\tilde{x}(t),\tilde{q})\right| - j^{-1} - 2\sigma(t)\left|j^{-1}\right|.$$

From (7.18), we conclude that

$$\left\langle \tilde{q}, \tilde{x}^{(\alpha)}(t)\right\rangle \geq \left|H(t,\tilde{x}(t),\tilde{q})\right| - j^{-1} - 2\sigma(t)\left|j^{-1}\right|.$$

Then, we have

$$\left|\left\langle \tilde{q}, \tilde{x}^{(\alpha)}(t)\right\rangle - H(t,\tilde{x}(t),\tilde{q})\right| \leq \left|j^{-1}\right|(2\sigma(t)+1). \tag{7.20}$$

As $j \to \infty$ in (7.20) yields condition (c) of Lemma 7.1. Since the condition (b) of the Lemma 7.1 implies $\tilde{q}(b)$ is bounded, by applying Theorem 6.2 of Chapter 6 of this thesis leads to $\tilde{p}(\cdot)$, $\gamma(\cdot)$, $\mu(\cdot)$, and ζ satisfying the required condition in the limit as $j \to \infty$.

Finally, we will show the extra hypothesis (IH2) can be deleted. For each j we defined

$$\Omega_j(t) = \left\{ v \in \Omega(t) : k(t,v) \leq k(t,u(t)) + j, \left| \tilde{f}(t,\tilde{x}(t),v) \right| \leq \left| \tilde{x}^{(\alpha)}(t) \right| + j \right\}.$$

Note that: $\Omega_j(t)$ is an increasing sequence of multifunctions, and that any element of $\Omega(t)$ belongs to $\Omega_j(t)$ for j sufficiently large. Now, we consider the problem obtained by replacing $\Omega(t)$ by $\Omega_j(t)$, and the hypotheses are satisfied. Then, we assume the existent of multipliers $\tilde{p}(\cdot)$, $\mu(\cdot)$, $\gamma(\cdot)$, and ζ all depending on j with the properties listed in Lemma 7.1 except that the maximization of the Hamiltonian condition now takes the form

$$\left\langle \tilde{q}, \tilde{x}^{(\alpha)} \right\rangle \geq \left\langle \tilde{q}, \tilde{f}(t,\tilde{x}(t),v) \right\rangle, \qquad \text{for all } v \in \Omega_j(t) \text{ a.e., } t \in [a,b]. \tag{7.21}$$

By applying Theorem 6.2 Chapter 6 yields $\tilde{p}(\cdot)$, $\mu(\cdot)$, $\gamma(\cdot)$, and ζ which continue to satisfy the conditions (a), (b) and (d) of the Lemma 7.1. By using the fact, for any $t \in [a,b]$, $v(t) \in \Omega(t)$ implies $v(t) \in \Omega_j(t)$ for j large enough. This fact with (7.21) satisfy the condition (c) of the Lemma 7.1 for the limiting data.

Then, we justifies all the assertions of Lemma 7.1. □

7.4 Illustrative Example

The FOCP with state constraints considered in this example to illustrate the application of the proved Maximum Principle of Pontryagin can be stated as follows.

$$\begin{aligned}
\text{Minimize} \quad & -y(T) & & \\
\text{subject to} \quad & x^{(\alpha)} = u(t)x(t), & x(0) = x_0, & \qquad (7.22)\\
& y^{(\alpha)} = (1-u(t))x(t), & y(0) = 0, & \qquad (7.23)\\
& u(t) \in [0,1], & & \qquad (7.24)\\
& x(t) \leq a + bt^{\alpha}, & & \qquad (7.25)
\end{aligned}$$

where all the relations hold Lebesgue a.e. in $[0,T]$ and the constants x_0, a, b and T satisfy $a > x_0 > 0$, $T > 1$, $b > 0$, and, there are constants c_1 and c_2 such that

$$c_2 \geq c_1,\ E_\alpha(c_1^\alpha) > \frac{a}{x_0},\ c_2 = T - (\Gamma(\alpha+1))^{\frac{1}{\alpha}}.$$

Notice that we are considering smooth data in order to facilitate the understanding of the issues involved when state constraints are present.

Let us denote optimal control process (x, y, u) (we omit asterisks) given by

$$u(t) = \begin{cases} 1 & \text{if } t \in [0, t_1) \\ \dfrac{b\,\Gamma(\alpha+1)}{a + bt^\alpha} & \text{if } t \in [t_1, t_2) \\ 0 & \text{if } t \in [t_2, T], \end{cases}$$

$$x(t) = \begin{cases} E_\alpha(t^\alpha)x_0 & \text{if } t \in [0, t_1) \\ a + bt^\alpha & \text{if } t \in [t_1, t_2) \\ a + bt_2^\alpha & \text{if } t \in [t_2, T], \end{cases}$$

$$y(t) = \begin{cases} 0 & \text{if } t \in [0, t_1) \\ \left(\dfrac{a}{\Gamma(\alpha+1)} - b\right)(t - t_1)^\alpha + \dfrac{b\Gamma(\alpha+1)}{\Gamma(2\alpha+1)}(t - t_1)^{2\alpha} & \text{if } t \in [t_1, t_2) \\ y(t_2) + \dfrac{a + bt_2^\alpha}{\Gamma(\alpha+1)}(t - t_2)^\alpha & \text{if } t \in [t_2, T], \end{cases}$$

where t_1 and t_2 - respectively, the first and the last times t for which $x(t) = a + bt^\alpha$ - exist due to the considered assumptions (for example, by considering $t_1 = c_1$ and $t_2 = c_2$) and are chosen in order to maximize the value of $y(T)$.

Next we show that the this control process satisfies the necessary conditions of optimality derived in this chapter, that is, $u(t) \in [0, 1]$ maximizes, Lebesgue a.e. on $[0, T]$, the map

$$v \to H(x(t), y(y), v, p(t), q(t)),$$

where the Pontryagin function H is defined by

$$H(x, y, u, p, q) = [(p + \gamma)u + q(1 - u)]\, x,$$

being $\gamma(t) = \dfrac{1}{\Gamma(\alpha+1)} \displaystyle\int_{[0,t)} (d\mu)^\alpha$ - and, by the way, $\gamma(t^+) = \dfrac{1}{\Gamma(\alpha+1)} \displaystyle\int_{[0,t]} (d\mu)^\alpha$ - is the fractional Stieltjes integral on the interval $[0, t)$ of the positive Radon Borel measure μ supported the set of points in $[0, T]$ on which the state constraint is active (i.e., $x(t) = a + bt^\alpha$), and (p, q) are the adjoint variables satisfying the following differential equations Lebesgue a.e. in $[0, T]$

$$\begin{aligned} &-q^{(\alpha)}(t) = 0, \qquad q(T) = 1 \\ &-p^{(\alpha)}(t) = (p(t) + \gamma(t))u(t) + q(t)(1 - u(t)), \qquad p(T) + \gamma(T^+) = 1. \end{aligned}$$

First, it is immediate to conclude that $q(t) = 1$, for all $t \in [0, T]$.

Now, for all $t \in [0, t_1)$, $u(t) = 1$ and, since $x(t) < a + bt^\alpha$, $\gamma(t) = 0$, and, from the fact that $-p^{(\alpha)}(t) = p(t)$, we have $p(t) = p(t_1)E_\alpha((t_1 - t)^\alpha)$. From the maximization of the Pontryagin function, it results that $p(t) \geq 1$, for $t \in [0, t_1)$, and, thus, $p(t_1) \geq 1$.

For all $t \in [t_1, t_2]$, we have $u(t) = \dfrac{b\Gamma(\alpha+1)}{a + bt^\alpha}$. It is not difficult to verify, that, under the above assumptions, $u(t) \in (0, 1)$. Moreover, from the maximization of the Pontryagin function, we conclude immediately that $p(t) + \gamma(t) = 1$. The fact that $-p^{(\alpha)}(t) = 1$ implies that $p(t) = p(t_1) - \dfrac{1}{\Gamma(\alpha+1)}(t - t_1)^\alpha$. Since

$$\gamma(t) = p(t) - p(t_1) + \frac{1}{\Gamma(\alpha+1)}(t - t_1)^\alpha \geq 0,$$

for all $t \in [t_1, t_2]$ implies not only that $p(t_1) = 1$, but also that $d\mu = dt$.

In the last time subinterval, $(t_2, T]$, we have that $u(t) = 0$, and from $-p^{(\alpha)}(t) = 1$, it follows that

$$p(t) = p(t_2) - \frac{1}{\Gamma(\alpha+1)}(t - t_2)^\alpha.$$

Since

$$p(t_2) = 1 - \frac{1}{\Gamma(\alpha+1)}(t_2 - t_1)^\alpha,$$

we conclude that

$$p(t) = 1 - \frac{1}{\Gamma(\alpha+1)} \left((t - t_2)^\alpha + (t_2 - t_1)^\alpha\right).$$

By using the fact that

$$p(T) = -\frac{1}{\Gamma(\alpha+1)} \int_{[t_1, t_2]} (dt)^\alpha = -\frac{1}{\Gamma(\alpha+1)}(t_2 - t_1)^\alpha,$$

we have that

$$t_2 = T - (\Gamma(\alpha+1))^{\frac{1}{\alpha}}.$$

From the continuity in time of the maximized Pontryagin function at $t = t_1$, i.e., $E_\alpha(t_1^\alpha)x_0 = a + bt_1^\alpha$, and the fact that

$$t_1^\alpha = \frac{a}{b}\left(\frac{x_0}{a}E_\alpha(t_1^\alpha) - 1\right) > 0,$$

we conclude that $E_\alpha(t_1^\alpha) > \dfrac{a}{x_0}$.

With some standard effort, it is not difficult to see that the values of t_1 and t_2 with $t_2 \geq t_1$ that maximize the value of $y(T)$, i.e., the ones that yield the minimum cost, are compatible with assumptions with $c_1 = t_1$ and $c_2 = t_2$ imposed on the data of the problem. Thus, we have shown that the optimal control process considered on an intuitive basis satisfies the maximum principle of Pontryagin proved in this chapter. Thus, it is not difficult to reconstruct the solution to the problem by choosing the controls that enforce the validity of the optimality conditions moving backwards from the final time.

Chapter 8

Conclusions and Prospective Research

8.1 Conclusions

The main contribution of this thesis has been introduced in Chapters 4, 5, 6 and 7. The other chapters in this dissertation deal with that, albeit the relevance in their own, they also played a subsidiary role in the formulation and derivation of the main results.

Our objective concerns the formulation of necessary conditions of optimality for fractional problems, where the characterization of the optimal-control problem in the fractional context is more accurate than for the integer counterpart. The fractional optimal-control problem can be viewed as a generalization for fractional calculus of the optimal-control problem in the integer sense.

We began by formulating the fractional optimal-control problem with data satisfying relatively strong assumptions. Subsequently, we increase the complexity by studying the problem in the absence of smoothness on its data. Also, the state constraints were imposed on the data of the problem.

In Chapter 4, under some smoothness assumptions, we derived a Pontryagin maximum principle for a general formulation of fractional optimal-control problems, whose cost function is in the fractional integral form and whose dynamics is characterized by the Caputo fractional derivative. Also, we presented a new technique to obtain a fractional-problem maximum principle, where the necessary conditions of optimality are derived with variations on the original problem, and not by converting the fractional problem to the classical one (with integer order) as done in the earlier literature. Furthermore, an illustrative example is solved by using the conditions of the maximum principle together with the Mittag-Leffler function, to show the effectiveness of the proposed approach.

In Chapter 5, we define the fractional integral with respect to a general Radon measure in the Jumarie sense, and we formulate this fractional integral in two cases: with and without atomic measure component. Besides, being new results in measure and integration theory for the fractional context, these results are especially relevant for Chapters 6 and 7.

In Chapter 6, we formulate and prove necessary conditions of optimality for fractional optimal-control problems with state constraints, being the dynamic system is modeled by a fractional differential inclusion in the Jumarie sense. Besides of the interest in their own, the results of this chapter are particularly helpful to obtain the maximum principle of Chapter 7.

In Chapter 7, a maximum principle for fractional optimal-control problem with state constraints, and with weak assumptions are presented and proved. The adopted approach follows the one in Clarke [43] is used in this chapter, to take advantage of the results obtained in Chapter 6. Moreover, the proposed approach is illustrated by an example.

To sum up: the main conclusion of this thesis consists in the fact that we extended in a number of very significant ways the current theory on necessary conditions of optimality so far developed for optimal fractional differential control problems as characterized in the contents of the various chapters of the thesis. In this way, this thesis constitutes a contribution to lessen the current existing gap between both bodies of theories.

8.2 Future Works

As it is clear from the wealth of issues that arise in optimal-control theory, there are a large number of issues that were left untouched due to the short period of time (3 years, including the scholar part) that were made available to devote my efforts to the proposed challenges. In this way, some subjects have not been explored, but are left for future works.

There are many points worth of future investigation, among which I would like to single out the following:

(1) In Chapters 6 and 7, only state constraints have been considered in the problem. In future works we can derive a maximum principle with mixed constraints under appropriate assumptions.

(2) It is not difficult to construct an example for which the proved maximum principles degenerate. An important issue that is of interest concerns the additional assumptions under which the maximum principle does not degenerate, that is the conditions remain informative.

(3) The results obtained in Chapter 5 open the door for the consideration of fractional impulsive dynamic control systems. We anticipate that the underlying inherent technical issues will be extremely challenging. However, given the range of applications, this is another direction along which the gap between integer and fractional optimal-control theories could become smaller.

(4) We recall that in this thesis we have only discussed the fractional derivative with respect to time t. However, there a large number of problems whose dynamics involve partial differential equations for which the fractional derivative with respect to the state constraints should be considered.

Appendices

Appendix A

Fractional Calculus

In this Appendix, we provide a brief review of some key concepts of fractional derivative and calculus. The fractional operators in Appendix A.2 are out of the scope of this thesis. Here, we just offer to the reader other types of fractional operators.

A.1 A Historical Review

The story of fractional calculus started when L'Hospital wrote to Leibniz a letter dated September 30th 1695, asking him what is the meaning of $\frac{d^n x}{dt^n}$, if $n = \frac{1}{2}$ (fractional), and Leibniz's response was: "An apparent paradox, from which one day useful consequences will be drawn ".

The question raised by Leibniz for a fractional derivative has been an ongoing topic in the last 300 years. Since then fractional calculus has attracted the attention of many famous mathematicians, such as Euler (1730), Lagrange (1772), Laplace (1812), Fourier (1822), Abel (1823–1826), Liouville (1832–1837), Riemann (1847), Grünwald (1867–1872), Letnikov (1868–1872), Heaviside (1892–1912), Weyl (1917), Erdélyi (1939–1965) and many others (see *e.g.*, Dalir and Bashour [47], and Gorenflo and Mainardi [62]). However, only since the Seventies fractional calculus has been the object of specialized conferences and treatises. The credit for the first open scientific event is due to B. Ross, who organized the first conference on fractional calculus and its applications at the university of New Haven in June 1974, and edited its proceedings. The first monograph devoted to fractional calculus was published in 1974 by K.B. Oldham and J. Spanier. It addresses their joint collaboration that began in 1968. This collaboration between a chemist (Oldham) and a mathematician (Spanier) in treating problems of mass and heat transfer in terms of the so-called semi-derivatives and semi-integrals clearly manifests the origin of a new era for fractional calculus, based on both physical intuition and mathematical versatility. In 1987, the large book by Samko, Kilbas and Marichev, referred to now as "encyclopedia" of fractional calculus, appeared first in Russian, and later (1993) translated into English. Nowadays, some series of books, journals and texts have been devoted to fractional calculus and its applications (Machado *et al.* [103]). This number is expected to grow in the forthcoming years.

A.2 Fractional Operators

A.2.1 Definitions of fractional integrals

Let $\alpha > 0$.

Definition A.1. Hadamard fractional integral

- *The left Hadamard fractional integral of order α is*

$$ {}_a^H I_t^\alpha f(t) = \frac{1}{\Gamma(\alpha)} \int_a^t (\ln \frac{t}{\tau})^{\alpha-1} \frac{f(\tau)}{\tau} d\tau, \qquad t \in [a,b]. $$

- *The right Hadamard fractional integral of order α is*

$$ {}_t^H I_b^\alpha f(t) = \frac{1}{\Gamma(\alpha)} \int_t^b (\ln \frac{\tau}{t})^{\alpha-1} \frac{f(\tau)}{\tau} d\tau, \qquad t \in [a,b]. $$

Definition A.2. Chen fractional integral

- *The left Chen fractional integral of order α is*

$$ I_c^\alpha f(t) = \frac{1}{\Gamma(\alpha)} \int_c^t (t-\tau)^{\alpha-1} f(\tau) d\tau, \qquad t > c. $$

- *The right Chen fractional integral of order α is*

$$ I_c^\alpha f(t) = \frac{1}{\Gamma(\alpha)} \int_t^c (\tau-t)^{\alpha-1} f(\tau) d\tau, \qquad t < c. $$

Definition A.3. Kober fractional integral

- *The left Kober fractional integral of order α is*

$$ I_{1,\eta}^\alpha f(t) = \frac{t^{-\alpha-\eta}}{\Gamma(\alpha)} \int_0^t (t-\tau)^{\alpha-1} f(\tau) d\tau. $$

- *The right Kober fractional integral of order α is*

$$ I_{1,\eta}^\alpha f(t) = \frac{t^{\eta}}{\Gamma(\alpha)} \int_t^\infty (\tau-t)^{\alpha-1} f(\tau) d\tau. $$

Definition A.4. Erdélyi fractional integral

- *The left Erdélyi fractional integral of order α is*

$$I^{\alpha}_{\sigma,\eta}f(t) = \frac{\sigma t^{-\sigma(\alpha+\eta)}}{\Gamma(\alpha)} \int_0^t (t^{\sigma} - \tau^{\sigma})^{\alpha-1} \tau^{\sigma\eta+\sigma-1} f(\tau) d\tau.$$

- *The right Erdélyi fractional integral of order α is*

$$I^{\alpha}_{\sigma,\eta}f(t) = \frac{\sigma t^{\sigma\alpha}}{\Gamma(\alpha)} \int_t^{\infty} (\tau^{\sigma} - t^{\sigma})^{\alpha-1} \tau^{\sigma(1-\alpha-\eta)-1} f(\tau) d\tau.$$

A.2.2 Definitions of fractional derivatives

Let $\alpha > 0$ and $n-1 < \alpha \leq n$, $n \in \mathbb{N}$.

Definition A.5. Liouville fractional derivative

- *The left Liouville fractional derivative of order α is*

$$D^{\alpha}_{+}f(t) = \frac{1}{\Gamma(n-\alpha)} \frac{d^n}{dt^n} \int_{-\infty}^t (t-\tau)^{n-\alpha-1} f(\tau) d\tau, \qquad t > 0.$$

- *The right Liouville fractional derivative of order α is*

$$D^{\alpha}_{-}f(t) = \frac{(-1)^n}{\Gamma(n-\alpha)} \frac{d^n}{dt^n} \int_t^{+\infty} (t-\tau)^{n-\alpha-1} f(\tau) d\tau, \qquad t < +\infty.$$

Definition A.6. Grünwald-Letnikov fractional derivative

- *The left Grünwald–Letnikov fractional derivative of order α is*

$${}^{GL}_{a}D^{\alpha}_{t}f(t) = \lim_{h\to 0} \frac{1}{h^{\alpha}} \sum_{k=0}^{[\frac{t-a}{h}]} (-1)^k \binom{\alpha}{k} f(t-kh).$$

- *The right Grünwald–Letnikov fractional derivative of order α is*

$${}^{GL}_{t}D^{\alpha}_{b}f(t) = \lim_{h\to 0} \frac{1}{h^{\alpha}} \sum_{k=0}^{[\frac{b-t}{h}]} (-1)^k \binom{\alpha}{k} f(t+kh).$$

Here, $\binom{\alpha}{k}$ is the generalization of binomial coefficients to real numbers, defined by

$$\binom{\alpha}{k} = \frac{\Gamma(\alpha+1)}{\Gamma(k+1)\Gamma(\alpha-k+1)}.$$

Definition A.7. Hadamard fractional derivative

- *The left Hadamard fractional derivative of order α is*

$$ {}_a^H D_t^\alpha f(t) = \frac{1}{\Gamma(n-\alpha)} \left(t\frac{d}{dt}\right)^n \int_a^t (\ln\frac{t}{\tau})^{n-\alpha-1}\frac{f(\tau)}{\tau}d\tau, \qquad t\in[a,b].$$

- *The right Hadamard fractional derivative of order α is*

$$ {}_t^H D_b^\alpha f(t) = \frac{1}{\Gamma(n-\alpha)} \left(-t\frac{d}{dt}\right)^n \int_t^b (\ln\frac{\tau}{t})^{n-\alpha-1}\frac{f(\tau)}{\tau}d\tau, \qquad t\in[a,b].$$

Definition A.8. Chen fractional derivative

- *The left Chen fractional derivative of order α is*

$$D_c^\alpha f(t) = \frac{1}{\Gamma(1-\alpha)}\frac{d}{dt}\int_c^t (t-\tau)^{-\alpha}f(\tau)d\tau, \qquad t>c.$$

- *The right Chen fractional derivative of order α is*

$$D_c^\alpha f(t) = -\frac{1}{\Gamma(1-\alpha)}\frac{d}{dt}\int_t^c (\tau-t)^{-\alpha}f(\tau)d\tau, \qquad t<c.$$

Definition A.9. Marchaud fractional derivative

- *The left Marchaud fractional derivative of order α is*

$$D_+^\alpha f(t) = \frac{\alpha}{\Gamma(1-\alpha)}\int_0^\infty \frac{f(t)-f(t-\tau)}{\tau^{\alpha+1}}d\tau.$$

- *The right Marchaud fractional derivative of order α is*

$$D_-^\alpha f(t) = \frac{\alpha}{\Gamma(1-\alpha)}\int_0^\infty \frac{f(t)-f(t+\tau)}{\tau^{\alpha+1}}d\tau.$$

Definition A.10. Riesz fractional derivative

$$D_t^\alpha f(t) = -\frac{1}{2\cos(\frac{\alpha\pi}{2})}\frac{1}{\Gamma(\alpha)}\frac{d^n}{dt^n}\left\{\int_{-\infty}^t (t-\tau)^{n-\alpha-1}f(\tau)d\tau + \int_t^\infty (\tau-t)^{n-\alpha-1}f(\tau)d\tau\right\}.$$

A.3 Relation Between the Fractional Derivatives

Here, we just recall some relations useful for our purposes (for their proofs see *e.g.*, Kilbas *et al.* [87], and Podlubny [125]).

The Riemann-Liouville and Caputo derivatives are related in the following way. Let $t > 0$, $\alpha \in \mathbb{R}$ and $n-1 < \alpha \leq n \in \mathbb{N}$. Then,

$$\begin{aligned}
{}_aD_t^\alpha f(t) &= {}_a^CD_t^\alpha f(t) + \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{\Gamma(k+1-\alpha)}(t-a)^{k-\alpha}, \\
{}_tD_b^\alpha f(t) &= {}_t^CD_b^\alpha f(t) + \sum_{k=0}^{n-1} \frac{f^{(k)}(b)}{\Gamma(k+1-\alpha)}(b-t)^{k-\alpha},
\end{aligned}$$

and

$$\begin{aligned}
{}_a^CD_t^\alpha f(t) &= {}_aD_t^\alpha \left(f(t) - \sum_{k=0}^{n-1} \frac{(t-a)^k}{\Gamma(k+1)} f^{(k)}(a) \right), \\
{}_t^CD_b^\alpha f(t) &= {}_tD_b^\alpha \left(f(t) - \sum_{k=0}^{n-1} \frac{(b-t)^k}{\Gamma(k+1)} f^{(k)}(b) \right).
\end{aligned}$$

A.4 Basic Properties of Fractional Calculus

Among the several properties of the operators of differentiation and integration of arbitrary order, here we express some of the most useful for our purposes, notably for the Riemann–Liouville and Caputo derivatives.

Proposition A.1. (The constant function).
For the Riemann–Liouville fractional derivative, for any constant k, we have

$$D^\alpha k = \frac{k}{\Gamma(1-\alpha)} t^{-\alpha}.$$

On the contrary, for the Caputo fractional derivative, for any constant k, we have

$${}^CD^\alpha k = 0.$$

Proposition A.2. (Linearity).
Let $n-1 < \alpha < n \in \mathbb{N}$, $f(t)$ and $g(t)$ two continuous functions defined on $[a, b]$ such that ${}_aD_t^\alpha f$ and ${}_aD_t^\alpha g$ exist almost everywhere. Moreover, let $\lambda_1, \lambda_2 \in \mathbb{R}$. Then, $D_a^\alpha(\lambda_1 f \pm \lambda_2 g)$ exists almost everywhere, and the Riemann–Liouville derivative obeys

$${}_aD_t^\alpha[\lambda_1 f(t) \pm \lambda_2 g(t)] = \lambda_1 {}_aD_t^\alpha f(t) \pm \lambda_2 {}_aD_t^\alpha g(t).$$

Similarly, the Caputo derivative satisfies

$$ {}_a^C D_t^\alpha [\lambda_1 f(t) \pm \lambda_2 g(t)] = \lambda_1 {}_a^C D_t^\alpha f(t) \pm \lambda_2 {}_a^C D_t^\alpha g(t). $$

Proposition A.3. (The semigroup property of the Riemann–Liouville integral operator). Let α, $\beta > 0$, $t > 0$, and $f(t) \in L^p(a,b)$, $1 \le p \le \infty$. Then,

$$ I^\alpha I^\beta f(t) = I^\beta I^\alpha f(t) = I^{\alpha+\beta} f(t), \quad t \in [a,b] \text{ a.e.}. $$

Proposition A.4. *Let $n-1 < \alpha \le n \in \mathbb{N}$, $t \in [a,b]$ and $f(t) \in L^p(a,b)$, $1 \le p \le \infty$. Then,*

$$ \begin{aligned} {}_a^C D_t^\alpha \, {}_a I_t^\alpha f(t) &= f(t), \\ {}_a I_t^\alpha \, {}_a^C D_t^\alpha f(t) &= f(t) - \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{\Gamma(k+1)} (t-a)^k. \end{aligned} $$

The same result holds when using the Riemann–Liouville derivative.

Proposition A.5. (Interpolation).
Let $n-1 < \alpha \le n \in \mathbb{N}$, $t \in [a,b]$, and $f(t)$ be a function such that $D^\alpha f(t)$ exists. Then, the Riemann–Liouville derivative obeys

$$ \begin{aligned} \lim_{\alpha \to n} {}_a D_t^\alpha f(t) &= f^{(n)}(t), \\ \lim_{\alpha \to n-1} {}_a D_t^\alpha f(t) &= f^{(n-1)}(t). \end{aligned} $$

For the Caputo derivative, the corresponding interpolation property reads:

$$ \begin{aligned} \lim_{\alpha \to n} {}_a^C D_t^\alpha f(t) &= f^{(n)}(t), \\ \lim_{\alpha \to n-1} {}_a^C D_t^\alpha f(t) &= f^{(n-1)}(t) - f^{(n-1)}(a). \end{aligned} $$

Proposition A.6. (Leibniz Rule).
Let $\alpha \in \mathbb{R}$, $t > 0$, $n-1 < \alpha \le n \in \mathbb{N}$ and $f(t), g(t)$ be continuous functions on $[a,b]$; then the generalized Leibniz formula for the Riemann–Liouville derivative is defined as

$$ D_{a+}^\alpha [f(t)g(t)] = \sum_{k=0}^{\infty} \binom{\alpha}{k} \left(D^{\alpha-k} f(t) \right) D^k [g(t)], $$

where we use the binomial coefficient

$$ \binom{\alpha}{k} = \frac{\Gamma(\alpha+1)}{\Gamma(k+1)\Gamma(\alpha-k+1)}. $$

The Leibniz formula for the Caputo derivative satisfies:

$$ {}_a^C D_t^{\alpha}[f(t)g(t)] = \sum_{k=0}^{\infty} \binom{\alpha}{k} \left(D^{\alpha-k} f(t) \right) D^k g(t) - \sum_{k=0}^{n-1} \frac{t^{k-\alpha}}{\Gamma(k-\alpha+1)} D^k[g(t)f(t)](a). $$

A.5 Generalized Taylor's Formula

Taylor's formula has been generalized by many authors (Odibat and Shawagfeh [116]). Here, we just recall only two formulas, Caputo and Riemann–Liouville.

- Generalization of Taylor's formula involving Caputo fractional derivatives.

  Let $0 < \alpha \leq 1$, $n \in \mathbb{N}$, and $f(x)$ a continuous function in $[a, b]$ (Odibat and Shawagfeh [116]). Then, for all $x \in [a, b]$, we have

  $$ f(x) = \sum_{k=0}^{n} \frac{(x-a)^{k\alpha}}{\Gamma(k\alpha+1)} \, {}_a^C D_x^{k\alpha} f(a) + R_n(x, a), $$

  where $R_n(x, a)$ is the remainder of the generalized Taylor's series defined by

  $$ R_n(x, a) = {}_a^C D_x^{(n+1)\alpha} f(\xi) \frac{(x-a)^{(n+1)\alpha}}{\Gamma((n+1)\alpha+1)}. $$

  Here, $a \leq \xi \leq x$ and ${}_a^C D_x^{\alpha}$ is the left Caputo fractional derivative of order α.

- Generalization of Taylor's formula in the Riemann–Liouville sense.

  Let $\alpha > 0$, $n \in \mathbb{Z}^+$, and $f(x) \in \mathbb{C}^{[\alpha]+n+1}([a, b])$ (Munkhammar [113]). Then,

  $$ f(x) = \sum_{k=-n}^{n-1} \frac{(x-x_0)^{k+\alpha}}{\Gamma(k+\alpha+1)} \, D_{a+}^{k+\alpha} f(x_0) + R_n(x), $$

  for all $a \leq x \leq b$, where $R_n(x)$ is the remainder defined by

  $$ R_n(x) = I_{a+}^{\alpha+n} D_{a+}^{\alpha+n} f(x). $$

  Here, D_{a+}^{α} is the left Riemann–Liouville fractional derivative of order α, I_{a+}^{α} is the left Riemann–Liouville fractional integral of order α, and $[\alpha]$ is the integer part of α.

A.6 Some Properties for Jumarie Fractional Derivative and Integral

Definition A.11. Jumarie fractional derivative via fractional difference.

Let $f : \mathbb{R} \to \mathbb{R}$, $x \to f(x)$ be a continuous function (but not necessarily differentiable), and $h > 0$ denote a constant discretization span. The forward operator $FW(h)$ is defined as:

$$FW(h).f(x) := f(x+h).$$

For $\alpha \in \mathbb{R}$ and $0 < \alpha \leq 1$, the fractional difference $\Delta^\alpha f(x)$ is defined by

$$\Delta^\alpha f(x) := (FW-1)^\alpha.f(x) = \sum_{k=0}^{\infty} (-1)^k \binom{\alpha}{k} f(x+(\alpha-k)h),$$

and the Jumarie fractional derivative of order α is

$$f^{(\alpha)}(x) = \lim_{h \downarrow 0} \frac{\Delta^\alpha [f(x) - f(0)]}{h^\alpha}.$$

The Jumarie fractional derivative has the following properties:

- The α^{th} derivatives of a constant is zero.
- Fractional Barrow's formula
$$\int_a^t f^{(\alpha)}(\tau)(d\tau)^\alpha = \Gamma(\alpha+1)(f(t) - f(a)).$$
- Fractional derivative of compounded functions
$$d^\alpha f \cong \Gamma(1+\alpha) df,$$
or, in term of fractional difference, $\Delta^\alpha f \cong \Gamma(1+\alpha)\Delta f$.
- Fractional Leibniz rule
$$(f(t)g(t))^{(\alpha)} = (f(t))^{(\alpha)} g(t) + (g(t))^{(\alpha)} f(t).$$
- Inverse of Mittag-Leffler function in Jumarie form
$$\int_0^x \frac{d^\alpha t}{t} = \ln_\alpha x, \quad x = E_\alpha(\ln_\alpha x).$$
- The fractional derivative via difference:
$$f^{(\alpha)}(t) = \lim_{h \downarrow 0} \frac{\Delta^\alpha f(t)}{h^\alpha} = \Gamma(1+\alpha) \lim_{h \downarrow 0} \frac{\Delta f(t)}{h^\alpha}, \quad 0 < \alpha \leq 1,$$
where $\Delta^\alpha f \cong \Gamma(1+\alpha)\Delta f$, and for the generalization form defined by
$$f^{(\alpha)}(t) = \Gamma(1+(\alpha-n)) \lim_{h \downarrow 0} \frac{\Delta f^{(n)}(t)}{h^{\alpha-n}}, \quad n < \alpha \leq n+1.$$

- The fractional derivative of a composition function (fractional Chain rule)

$$\begin{aligned} f^{(\alpha)}[x(t)] &= \frac{df(x)}{dx} x^{(\alpha)}(t), \\ &= f_x^{(\alpha)}(x)(\frac{dx(t)}{dt})^{\alpha} = \Gamma(2-\alpha)x^{\alpha-1}f_x^{(\alpha)}(x)x^{(\alpha)}(t). \end{aligned}$$

Note that: In the formula of the fractional derivative of a composition function, $x(t)$ is non-differentiable in the first equation and differentiable in the second one, but $f(x)$ is differentiable in the first equation and non-differentiable in the second one.

- A generalization of Taylor's expansion

Proposition A.7. *Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous function, $x \to f(x)$ have fractional derivative of order $k\alpha$, for any positive integer k, and $0 < \alpha \leq 1$. Then, the fractional Taylor series is given by:*

$$f(x+h) = \sum_{k=0}^{\infty} \frac{h^{\alpha k}}{\Gamma(1+\alpha k)} f^{(\alpha k)}(x), \quad 0 < \alpha \leq 1.$$

(For the proof and further details, see Jumarie [76].) Moreover, this series can be written as

$$f(x+h) = E_{\alpha}(h^{\alpha} D_x^{\alpha}) f(x),$$

where $E_{\alpha}(\cdot)$ denotes the Mittag-Leffler function defined in Chapter 2.

Note that: This fractional Taylor series only applies on the non-differentiable functions. So, it does not work with the standard Riemann–Liouville derivative (Jumarie [77]).

Corollary A.1. *Assume that $n < \alpha \leq n+1$, $n \in \mathbb{N} - \{0\}$, and that $f(x)$ has derivatives of order k (integer), $1 \leq k \leq n$, and assume that $f^{(n)}(x)$ has a fractional Taylor's series of order $\alpha - n =: \beta$, provided by the expression*

$$f^{(n)}(x+h) = \sum_{k=0}^{\infty} \frac{h^{k(\alpha-n)}}{\Gamma(1+k(\alpha-n))} D^{k(\alpha-n)} f^{(n)}(x), \quad n < \alpha \leq n+1.$$

Then, integrating this series with respect to h, we have

$$f(x+h) = \sum_{k=0}^{n} \frac{h^k}{k!} f^{(k)}(x) + \sum_{k=1}^{\infty} \frac{h^{(k\beta+n)}}{\Gamma(1+(k\beta+n+1))} f^{(k\beta+n)}(x), \quad \beta := \alpha - n.$$

Theorem A.1. (Golbabai and Sayevand [61]). Assume that $f(x)$ is a continuous function and has fractional derivative of order α, then for $0 < \alpha \leq 1$, we have

$$\frac{d^{\alpha}}{dx^{\alpha}} J^{\alpha} f(x) = f(x),$$

$$J^{\alpha}\frac{d^{\alpha}}{dx^{\alpha}}f(x) \quad = \quad f(x) - f(0).$$

Proof. Form the definition of the fractional integration, we have

$$\begin{aligned}
\frac{d^{\alpha}}{dx^{\alpha}}J^{\alpha}f(x) &= \frac{d^{\alpha}}{dx^{\alpha}}\left(\frac{1}{\Gamma(\alpha)}\int_0^x (x-\tau)^{\alpha-1}f(\tau)d\tau\right) \\
&= \frac{d^{\alpha}}{dx^{\alpha}}\left(\frac{1}{\Gamma(\alpha)}\frac{1}{\alpha}\int_0^x f(\tau)(d\tau)^{\alpha}\right) \\
&= \frac{1}{\Gamma(\alpha+1)}\frac{d^{\alpha}}{dx^{\alpha}}\left(\int_0^x f(\tau)(d\tau)^{\alpha}\right) = \frac{1}{\Gamma(\alpha+1)}\Gamma(\alpha+1)f(x) \\
&= f(x).
\end{aligned}$$

Similarly,

$$\begin{aligned}
J^{\alpha}\frac{d^{\alpha}}{dx^{\alpha}}f(x) &= \frac{1}{\Gamma(\alpha+1)}\left(\int_0^x \left(\frac{d^{\alpha}}{d\tau^{\alpha}}f(\tau)\right)(d\tau)^{\alpha}\right) \\
&= \frac{1}{\Gamma(\alpha+1)}\Gamma(\alpha+1)f(\tau)\mid_{\tau=0}^{\tau=x} \\
&= f(x) - f(0).
\end{aligned}$$

$\square$

Appendix B

Variational Results

B.1 Exact Penalization

The idea is to transform constrained problems into unconstrained ones by adding to the original objective function a term that penalizes any violation of the constraint.

The next theorem gives the conditions under which a minimizer for a constrained optimization problem is also a minimizer for an unconstrained problem when the data are Lipschitz conditions.

Theorem B.1. (Exact Penalization Theorem (Vinter [150])).
Let (X, M) be a metric space, $C \subset X$ be a set, and $f\colon X \to R$ be a function Lipschitz continuous on X with Lipschitz constant K. Suppose that the point $\tilde{x}$ is a minimizer for the constrained minimization problem

$$\begin{aligned} &\text{Minimize} \quad f(x) \quad \text{over} \quad x \in R^k, \\ &\text{satisfying} \quad x \in C. \end{aligned}$$

Then, for any $\tilde{K} \geq K$, the point $\tilde{x}$ is a minimizer also for the unconstrained problem

$$\begin{aligned} \text{Minimize} &\quad f(x) + \tilde{K} d_C(x), \\ \text{over points} &\quad x \in R^k, \end{aligned}$$

where $d_C(x)$ is a distance function on X defined as

$$d_C(x) := \inf_{x' \in C} M(x, x'), \qquad \text{for each } x \in X.$$

Note that, if $\tilde{K} > K$ and C is a closed set, then the converse assertion is also true, *i.e.*, any minimizer $\tilde{x}$ for the unconstrained problem is also a minimizer for the constrained problem, in particular $\tilde{x} \in C$.

B.2 Ekeland Theorem

The idea of this theorem is that, if a point u approximately minimizes a function $f(\cdot)$, then some neighboring point $\bar{u}$ close to u is a minimizer for some new perturbed function $\bar{f}(\cdot)$, obtained by adding a small perturbation term to the original function $f(\cdot)$.

Theorem B.2. (Ekeland's Theorem (Ekeland [55])).
Let (V, Δ) be a complete metric space, $F\colon V \to \mathbb{R} \cup \{+\infty\}$ be a lower semi-continuous function bounded from below, and $u \in V$ a point. If u is almost a minimizer for V, satisfying

$$F(u) \leq \inf F + \varepsilon,$$

for some $\varepsilon \succ 0$, then for every $\lambda \succ 0$ there exists a nearby point $v \in V$ which is an actual minimizer for a slightly perturbed function, such that

1. $F(v) \leq F(u)$,
2. $\Delta(u, v) \leq \lambda$,
3. $F(v) < F(w) + \frac{\varepsilon}{\lambda}\Delta(w, v), \quad \forall w \neq v.$

Proof. Before we prove this theorem, we present the next lemma.

Lemma B.1. *Let S be a closed subset of $V \times R$, such that, for some scalar m, every element $(v, r) \in S$ satisfies $r \geq m$. Then, for every $(v_1, r_1) \in S$, there exists an element $(\bar{v}, \bar{r}) \in S$ satisfying $(v_1, r_1) \leq_\alpha (v, \bar{r})$ which is maximal in S for the partial order $\leq_\alpha$.*

Definition B.1. *Here, we will define partial ordering (Bishop and Phelps [32]). For any $\alpha > 0$, the partial ordering $\leq_\alpha$ on $V \times R$ is defined by:*

$$(v_1, r_1) \leq_\alpha (v_2, r_2) \Leftrightarrow r_2 - r_1 + \alpha\Delta(v_1, v_2) \leq 0.$$

This relation is reflexive, antisymmetric and transitive.

Now, we will prove that partial ordering satisfies all these relations.

First, $\leq_\alpha$ is reflexive. Let $(v_1, r_1) \in V \times R$, such that $(v_1, r_1) \leq_\alpha (v_1, r_1)$. Then,

$$r_1 - r_1 + \alpha\Delta(v_1, v_1) = \alpha\Delta(v_1, v_1) \leq 0,$$

where $\Delta(v_1, v_1) = 0$ because Δ is a distance. Then, $(v_1, r_1) \leq_\alpha (v_1, r_1)$. The first relation is proven.

Second, $\leq_\alpha$ is antisymmetric. Let $(v_1, r_1), (v_2, r_2) \in V \times R$, we know that

$$\begin{aligned}
(v_1, r_1) \quad &\leq_\alpha \quad (v_2, r_2) \Leftrightarrow r_2 - r_1 + \alpha\Delta(v_1, v_2) \leq 0, \\
(v_2, r_2) \quad &\leq_\alpha \quad (v_1, r_1) \Leftrightarrow r_1 - r_2 + \alpha\Delta(v_2, v_1) \leq 0.
\end{aligned}$$

Then,

$$(r_2 - r_1 + \alpha\Delta(v_1, v_2)) + (r_1 - r_2 + \alpha\Delta(v_2, v_1)) \leq 0.$$

This means that $2\alpha\Delta(v_1, v_2) \leq 0$ and, from the definition $\alpha > 0$, then $\Delta(v_1, v_2) \leq 0$. So,

$$\Delta(v_1, v_2) = 0 \Leftrightarrow v_1 = v_2.$$

After substituting $\Delta(v_1, v_2) = 0$, we have

$$\begin{aligned} r_2 - r_1 \leq 0 \quad &\Leftrightarrow \quad r_2 \leq r_1, \\ r_1 - r_2 \leq 0 \quad &\Leftrightarrow \quad r_1 \leq r_2. \end{aligned}$$

This means that $r_1 = r_2$, then $(v_1, r_1) = (v_2, r_2)$. The second relation is proven.

Third, $\leq_\alpha$ is transitive. Let $(v_1, r_1), (v_2, r_2), (v_3, r_3) \in V \times R$, such that

$$\begin{aligned} (v_1, r_1) \leq_\alpha (v_2, r_2) \quad &\Leftrightarrow \quad r_2 - r_1 + \alpha\Delta(v_1, v_2) \leq 0, \\ (v_2, r_2) \leq_\alpha (v_3, r_3) \quad &\Leftrightarrow \quad r_3 - r_2 + \alpha\Delta(v_2, v_3) \leq 0. \end{aligned}$$

So, we need to prove that $(v_1, r_1) \leq_\alpha (v_3, r_3)$. Computing $r_3 - r_1 + \alpha\Delta(v_1, v_3)$ by adding and removing r_2, and using the triangle inequality, we have

$$\begin{aligned} r_3 - r_1 + \alpha\Delta(v_1, v_3) \quad &= \quad r_3 - r_2 + r_2 - r_1 + \alpha\Delta(v_2, v_3) \\ &\leq \quad r_3 - r_2 + r_2 - r_1 + \alpha(\Delta(v_1, v_2) + \Delta(v_2, v_3)), \end{aligned}$$

therefore,

$$(r_3 - r_2 + \alpha\Delta(v_2, v_3)) + (r_2 - r_1 + \alpha\Delta(v_1, v_2)) \leq 0.$$

Then, $(v_1, r_1) \leq_\alpha (v_3, r_3)$. The third relation is proven.

Now, we come back to the proof of the lemma. Let S be a closed subset of $V \times R$, such that, for some scalar m, every element $(v, r) \in S$ satisfies $r \geq m$.

Let (v_n, r_n) be a sequence in S, (v_1, r_1) be the first element in this sequence, and (v_n, r_n) be known. Then,

$$\begin{aligned} S_n \quad &:= \quad \{(v, r) \in S : (v_n, r_n) \leq_\alpha (v, r)\}, &\text{(B.1)} \\ m_n \quad &:= \quad \inf\{r : (v, r) \in S_n \quad \text{for some } v \in V\}. &\text{(B.2)} \end{aligned}$$

By the lemma, for all element of S, we have $r \geq m$, so $S_n \leq S$ and $m_n \geq m$.

Let (v_{n+1}, r_{n+1}) be any point in S_n, such that

$$r_n - r_{n+1} \geq \frac{1}{2}(r_n - m_n). \tag{B.3}$$

S_n are closed and nested:

$$\begin{aligned} s_n &= \{(v,r) \in S : r - r_n + \alpha\Delta(v_n, v) \leq 0\}, \\ &= \{(v,r) \in S : \Delta(v_n, v) \leq \frac{r_n - r}{\alpha}\}. \end{aligned}$$

Similarly, we can define S_{n+1} as the following

$$\begin{aligned} s_{n+1} &= \{(v,r) \in S : r - r_{n+1} + \alpha\Delta(v_{n+1}, v) \leq 0\}, \\ &= \{(v,r) \in S : \Delta(v_{n+1}, v) \leq \frac{r_{n+1} - r}{\alpha}\}, \end{aligned}$$

but $(v_n, r_n) \leq_\alpha (v_{n+1}, r_{n+1})$, because $(v_{n+1}, r_{n+1}) \in S_n$. Since $\leq_\alpha$ is transitive as proven before, then $(v_n, r_n) \leq_\alpha (v, r) \Rightarrow (v, r) \in S_n$ and $(v, r) \in S_{n+1}$, then $S_{n+1} \subset S_n$.

From the definition of S_n, for which $(v_{n+1}, r_{n+1}) \in S_n$, $r_{n+1} \leq r_n$, and from (B.3) we have:

$$\begin{aligned} |r_{n+1} - m_{n+1}| &= |r_{n+1} - r_n + r_n - m_n + m_n - m_{n+1}| \\ &\leq \left|\frac{1}{2}(m_n - r_n) + r_n - m_n\right| = \frac{1}{2}\left|r_n - m_n\right| \\ &\leq \frac{1}{2}(\frac{1}{2}\left|r_{n-1} - m_{n-1})\right| \leq \frac{1}{2^2}\left|r_{n-1} - m_{n-1})\right| \leq \cdots \\ &\leq \frac{1}{2^n}\left|r_1 - m_1)\right| \leq \frac{1}{2^n}\left|r_1 - m)\right|. \end{aligned}$$

Hence, for every $(v, r) \in S_{n+1}$, we have

$$(v_{n+1}, r_{n+1}) \leq_\alpha (v, r) \Leftrightarrow r - r_{n+1} + \alpha\Delta(v_{n+1}, v) \leq 0, \quad m_{n+1} \leq r.$$

From ((B.1),(B.2)), we get

$$|r_{n+1} - r| \leq |r_{n+1} - m_{n+1}| \leq \frac{1}{2^n}\left|r_1 - m\right|.$$

Here, $r - r_{n+1} + \alpha\Delta(v_{n+1}, v) \leq 0 \Rightarrow \Delta(v_{n+1}, v) \leq \frac{1}{\alpha}\left|r_{n+1} - r\right|$, then

$$|\Delta(v_{n+1}, v)| = \Delta(v_{n+1}, v) \leq \frac{1}{\alpha}\left|r_{n+1} - r\right| \leq \frac{1}{2^n\alpha}\left|r_1 - m\right|.$$

Thus,

$$\lim_{n\to\infty} \frac{1}{2^n\alpha}\left|r_1 - m\right| = \frac{1}{2^\infty\alpha}\left|r_1 - m\right| = 0.$$

So, $0 \leq \Delta(v_{n+1}, v) \leq \frac{1}{2^n\alpha}\left|r_1 - m\right| \to 0$. This shows that the diameter of S_n goes to zero as $n \to \infty$. Since $V \times R$ has a complete metric, the sets S_n have one point $(\bar{v}, \bar{r})$ in common:

$$(\bar{v}, \bar{r}) = \bigcap_{n\geq 1} S_n.$$

By the definition of partial order, $(v_n, r_n) \leq_\alpha (\bar{v}, \bar{r})$ for every n. Suppose that $(\tilde{v}, \tilde{r}) \in S$, such that $(\bar{v}, \bar{r}) \leq_\alpha (\tilde{v}, \tilde{r}), \quad \forall n \in \mathbb{N}$, then, by the transitivity of partial ordering, we have $(v_n, r_n) \leq_\alpha \ (\tilde{v}, \tilde{r}), \quad \forall n \in \mathbb{N}$, whence $(\tilde{v}, \tilde{r}), \in \bigcap_{n \geq 1} S_n$, and therefore $(\bar{v}, \bar{r}) = (\tilde{v}, \tilde{r})$. This means that the element $(\bar{v}, \bar{r})$ is the maximal in S.

To prove the theorem, let

$$S = epi(V) = \{(v, F(v)) : v \in V, F(v) \in \mathbb{R}\},$$

and apply in the previous lemma $\alpha = \frac{\varepsilon}{\lambda}$ and $(v_1, r_1) = (u, F(u))$. Then, for the maximal element $(v, r) \in S$ satisfying

$$(u, F(u)) \leq_\alpha (v, r), \tag{B.4}$$

since (v, r) lies in S, we have $(v, r) \leq_\alpha (v, F(v)) \Rightarrow r = f(v)$. By the maximality of (v, r), expression (B.4) becomes

$$(u, F(u)) \leq_\alpha (v, F(v)) \Leftrightarrow F(v) - F(u) + \alpha\Delta(u, v) \leq 0, \tag{B.5}$$

since $\alpha\Delta(u, v) \geq 0$, then $F(v) - F(u) \leq 0 \Leftrightarrow F(v) \leq F(u)$, which gives the first condition of the theorem. The maximality of $(v, F(v)) \in S$ implies that, for any $w \in V$ such that $w \neq v$, $F(w)$ is finite, then the relation $(v, F(v)) \leq_\alpha (w, F(w))$ does not hold. So

$$F(w) - F(v) + (\frac{\varepsilon}{\lambda})\Delta(v, w) > 0 \Leftrightarrow F(w) + (\frac{\varepsilon}{\lambda})\Delta(v, w) > F(v),$$

which means the third condition of the theorem is proven.

Finally, because $F(u) \leq \inf(F) + \varepsilon$, then there exists $F(v) \geq F(u) - \varepsilon$, and by combining this relation with (B.5) we have:

$$\begin{aligned} F(v) - F(u) + \alpha\Delta(u, v) \leq 0 \quad &\Leftrightarrow \quad F(u) - \varepsilon - F(u) + \alpha\Delta(u, v) \leq 0 \\ &\Leftrightarrow \quad \alpha\Delta(u, v) \leq \varepsilon \Rightarrow \Delta(u, v) \leq \frac{\varepsilon}{\alpha} = \lambda. \end{aligned}$$

Then, the second condition of the theorem is proven. □

B.3 Generalized Gronwall Inequality

The Gronwall inequality has an important role in numerous differential and integral equations. The classical form of this inequality is described by the following theorem (Corduneanu [46]).

Theorem B.3. *For any $t \in [t_0, T]$, let $a(t)$, $b(t)$ and $w(t)$ be continuous functions, with $b(t) \geq 0$. If $w(t)$ satisfies*

$$w(t) \leq a(t) + \int_{t_0}^{t} b(\tau)w(\tau)d\tau,$$

where $b(t) \geq 0$, *then*

$$w(t) \leq a(t) + \int_{t_0}^{t} a(\tau)b(\tau)\exp\left(\int_{\tau}^{t} b(s)ds\right)d\tau.$$

In particular, if $a(t)$ *is non-decreasing, then*

$$w(t) \leq a(t)\exp\left(\int_{t_0}^{t} b(\tau)d\tau\right).$$

Now, we present a generalization of the Gronwall inequality which can be used in a fractional differential equation. There are several generalizations of the Gronwall–Bellman inequalities (see *e.g.*, Lin [95], Ye *et al.* [153], and Zheng [157]), let us recall the following one.

Theorem B.4. *Let* $\alpha > 0$, $a(t)$ *be a non-negative function locally integrable on* $t \in [0, T]$ *(where* $T \leq +\infty$*), and* $b(t)$ *be a non-negative, non-decreasing continuous function defined on* $0 \leq t \leq T$*, where* $b(t)$ *is bounded by a positive constant* K *(*i.e.*,* $b(t) \leq K$*). If* $w(t)$ *is non-negative and locally integrable on* $t \in [0, T]$*, and satisfies*

$$w(t) \leq a(t) + b(t)\int_{0}^{t} (t-\tau)^{\alpha-1}w(\tau)d\tau, \tag{B.6}$$

then,

$$w(t) \leq a(t) + \int_{0}^{t}\left[\sum_{n=1}^{\infty}\frac{(b(t)\Gamma(\alpha))^n}{\Gamma(n\alpha)}(t-\tau)^{n\alpha-1}a(\tau)\right]d\tau. \tag{B.7}$$

Proof. Let $\theta(t)$ be a locally integrable function, and let us define an operator B on θ as follows:

$$B\theta(t) := b(t)\int_{0}^{t}(t-\tau)^{\alpha-1}\theta(\tau)d\tau, \quad t \geq 0.$$

From inequality (B.6), we have

$$w(t) \leq a(t) + Bw(t),$$

this implies

$$w(t) \leq \sum_{k=0}^{n-1} B^k a(t) + B^n w(t). \tag{B.8}$$

In order to get the desired inequality, using (B.7) and (B.8), we prove that

$$B^n w(t) \leq \int_{0}^{t}\frac{(b(t)\Gamma(\alpha))^n}{\Gamma(n\alpha)}(t-\tau)^{n\alpha-1}w(\tau)d\tau, \tag{B.9}$$

and $B^n w(t)$ vanishes as n increases (*i.e.*, $B^n w(t) \to 0$ as $n \to +\infty$) for each $t \in [0, T)$.

We will use the mathematical induction method to verify the inequality in (B.9). First, we know that the inequality in (B.9) is true for $n = 1$. Second, we assume that the inequality in (B.9) is true for $n = k$, then we prove that it is also true for $n = k + 1$:

$$B^{k+1}w(t) = B(B^k w(t)) \leq b(t)\int_0^t (t-\tau)^{\alpha-1}\left[\int_0^\tau \frac{(b(s)\Gamma(\alpha))^k}{\Gamma(k\alpha)}(t-s)^{k\alpha-1}w(s)ds\right]d\tau.$$

Since $b(t)$ is a non-negative and non-decreasing function, it follows that

$$B^{k+1}w(t) \leq b^{k+1}(t)\int_0^t (t-\tau)^{\alpha-1}\left[\int_0^\tau \frac{(\Gamma(\alpha))^k}{\Gamma(k\alpha)}(t-s)^{k\alpha-1}w(s)ds\right]d\tau,$$

by interchanging the order of integration, we have:

$$B^{k+1}w(t) \leq b^{k+1}(t)\int_0^t \left[\int_s^t \frac{(\Gamma(\alpha))^k}{\Gamma(k\alpha)}(t-\tau)^{\alpha-1}(\tau-s)^{k\alpha-1}d\tau\right]w(s)ds.$$

Making the substitution $\tau = s + z(t-s)$ in the previous integral, and using the definition of Beta function (see *e.g*, Podlubny [125]), we obtain:

$$\begin{aligned}
\int_s^t (t-\tau)^{\alpha-1}(\tau-s)^{k\alpha-1}d\tau &= (t-s)^{k\alpha+\alpha-1}\int_0^1 (1-z)^{\alpha-1}z^{k\alpha-1}dz \\
&= (t-s)^{(k+1)\alpha-1}B(k\alpha,\alpha) \\
&= \frac{\Gamma(\alpha)\Gamma(k\alpha)}{\Gamma((k+1)\alpha)}(t-s)^{(k+1)\alpha-1}.
\end{aligned}$$

Then, we have

$$B^{k+1}w(t) \leq \int_0^t \frac{(b(t)\Gamma(\alpha))^{k+1}}{\Gamma((k+1)\alpha)}(t-s)^{(k+1)\alpha-1}w(s)ds,$$

and the inequality (B.9) is proved.

Since $B^n w(T) \leq \int_0^t \frac{(K\Gamma(\alpha))^n}{\Gamma(n\alpha)}(t-\tau)^{n\alpha-1}w(\tau)d\tau \to 0$, as $n \to +\infty$, for all $t \in [0,T)$, the proof is completed. $\square$

Corollary B.1. *Suppose the hypotheses presented in Theorem (B.4) are satisfied, and let $a(t)$ be a non-decreasing on $t \in [0,T]$. Then,*

$$\begin{aligned}
w(t) &\leq a(t)\left[1 + \int_0^t \sum_{n=1}^{\infty} \frac{(b(t)\Gamma(\alpha))^n}{\Gamma(n\alpha)}(t-\tau)^{n\alpha-1}d\tau\right] \\
&\leq a(t)E_\alpha(b(t)\Gamma(\alpha)t^\alpha),
\end{aligned}$$

where $E_\alpha(\cdot)$ is the generalized Mittag-Leffler function and $\Gamma(\cdot)$ is the gamma function.

Proof. From the Proof of Theorem (B.4), we have

$$w(t) \leq a(t) + \int_0^t \left[\sum_{n=1}^{\infty} \frac{(b(t)\Gamma(\alpha))^n}{\Gamma(n\alpha)} (t-\tau)^{n\alpha-1} a(\tau) \right] d\tau.$$

Since $a(t)$ is non-decreasing, we can write

$$\begin{aligned} w(t) &\leq a(t) \left[1 + \int_0^t \sum_{n=1}^{\infty} \frac{(b(t)\Gamma(\alpha))^n}{\Gamma(n\alpha)} (t-\tau)^{n\alpha-1} d\tau \right] \\ &\leq a(t) \sum_{n=0}^{\infty} \frac{(b(t)\Gamma(\alpha))^n}{\Gamma(n\alpha+1)} t^{n\alpha} \\ &\leq a(t) E_\alpha(b(t)\Gamma(\alpha) t^\alpha). \end{aligned}$$

The Corollary (B.1) is proved. □

B.4 Fractional DuBois-Reymond Fundamental Lemma

There are several forms of the fractional DuBois-Reymond Lemma (see *e.g.*, Almeida and Torres [9], Bourdin and Idczak [33], Kamocki [82], and Lazo and Torres [93]), here we will introduce the ones most important for our work.

Lemma B.2. (Lazo and Torres [93]). Let h be a differentiable function in the interval $[a, b]$ with $h(a) = 0$, $h(b) = 0$, and let $f \in L_1([a, b])$ be such that there is a number $\delta \in [a, b]$ with $|f(t)| \leq k(x-a)^\beta$ for all $t \in [a, \delta]$, where $k > 0$ and $\beta > -\alpha$ are constants. Then,

$${}_aI_b^\alpha(f(t)\,{}_aD_t^\alpha h(t)) = 0,$$

and

$$f(t) = c,$$

where c is a constant, $D^\alpha(\cdot)$ is the Riemann-Liouville fractional derivative operator, and I^α is the fractional integral operator.

Lemma B.3. (Almeida and Torres [9]). Let f be a continuous function satisfying

$$\int_a^b f(t)g(t)(dt)^\alpha = 0,$$

for every continuous function g satisfying $g(a) = g(b) = 0$. Then, $f = 0$. Here, the integral $\int_a^b (\cdot)(dt)^\alpha$ is the Jumarie fractional integral operator.

Appendix C

Nonsmooth Analysis

Nonsmooth analysis is an important tool in optimal-control theory. It first appeared when the classical analysis failed to give estimated approximations to non-differentiable functions and to the sets with non-differentiable boundaries, where there are many functions which are continuous everywhere but not differentiable at some points.

The first idea to develop nonsmooth analysis was taken from the geometric relationship between the derivative of smooth functions (differentiable functions) and the graph of these functions as follows.

In classical smooth analysis, the derivatives of a function g are related to vectors normal to tangent hyperplanes; for any differentiable function g the vector $(g'(x), -1)$ is a downward normal to the graph of g at $(x, g(x))$. Here, the graph of g is defined by

$$\operatorname{Gr} g = \{(x, r) \in \mathbb{R}^n \times \mathbb{R} \colon r = g(x)\}.$$

Instead of considering derivatives as elements of normal subspaces to smooth sets, generalized derivatives are defined to be elements of normal cones to possibly nonsmooth sets.

To tackle optimal-control issues, we use the following results.

Definition C.1. *Let X be a subset of a Banach space ψ. A function $g\colon X \to \mathbb{R}$ is said to satisfy a Lipschitz condition on X if*

$$|g(x_1) - g(x_2)| \leq k \|x_1 - x_2\|,$$

for all points $x_1, x_2 \in X$, where k is a positive constant, also referred to as Lipschitz condition of rank k. We say that g is Lipschitz of rank k near x if, for some $\delta > 0$, g satisfies a Lipschitz condition of rank k on the set $x + \delta B$, where B is the open unit ball.

Proposition C.1. *Let C be a nonempty subset of X, and $d_C(\cdot)\colon X \to \mathbb{R}$ be a distance function defined by*

$$d_C(x) = \inf\{\|x - c\| : c \in C\}.$$

Then, the distance function $d_C(\cdot)$ satisfies the Lipschitz condition on X as follows:

$$|d_C(x) - d_C(y)| \leq \|x - y\|.$$

Theorem C.1. *(Clarke [43]). Let $h(\cdot)$ be Lipschitz near x, suppose S is any set of Lebesgue measure 0 in $\mathbb{R}^n$, and let Ω_h be the set of points where a given function $h(\cdot)$ fails to be differentiable. Then, the generalized gradient is defined as:*

$$\bar{\partial}_x h(x) := co\{\lim \nabla h(x_i) : x_i \to x, x_i \notin S, x_i \notin \Omega_h\}.$$

Definition C.2. *Let x be a point in the set $C \subseteq X$, and $\lambda \in X$ be a tangent vector to C at x satisfying $d_C(x; \lambda) = 0$. Then, the set of all the tangents to C at x is called the tangent cone $T_C(x)$ and is defined by*

$$T_C(x) := \{\lambda \in \mathbb{R}^n \colon d_C(x; \lambda) = 0\},$$

where $d_C(x; \lambda)$ is the directional derivative of the distance function.

Definition C.3. *Let x be a point in $C \subseteq X$, and $\lambda \in X$ be a tangent vector to C at x. Then, the normal cone to C at x, $N_C(x)$, is defined as:*

$$N_C(x) := \{\xi \in X^* \colon \langle \xi, \lambda \leq 0 \rangle \quad \forall \lambda \in T_C(x)\}.$$

The limiting normal cone was introduced by Mordukhovich in [110, 111], as follows:

Let C be a nonempty subset of $\mathbb{R}^n$, and let

$$P(x, C) := \{z \in \operatorname{cl} C : \|x - z\| = d(x, C)\}$$

be the set of best approximations of x in $\operatorname{cl} C$ with respect the Euclidean distance function $d(x, C)$.

Definition C.4. *Given $\bar{x} \in \operatorname{cl} C$ the following closed cone*

$$N(\bar{x}, C) := \limsup_{x \to \bar{x}} \{\operatorname{cone}(x - P(x, C))\}$$

is called the normal cone to the set C a the point $\bar{x}$. If $\bar{x} \notin \operatorname{cl} C$, we put $N(\bar{x}, C) = \emptyset$.

Proposition C.2. *The normal cone $N_C(x)$ is the closed convex cone generated by $\partial d_C(x)$ and satisfies*

$$N_C(x) = \operatorname{cl}\{\bigcup_{\lambda \geq 0} \lambda \partial d_C(x)\},$$

where cl *is weak* closure.*

Corollary C.1. *Let $X = X_1 \times X_2$, where X_1, X_2 are Banach spaces, $C = C_1 \times C_2$, where C_1 and C_2 are subsets of X_1 and X_2, respectively. Suppose that $x = (x_1, x_2) \in C$. Then,*

$$T_C(x) = T_{C_1}(x_1) \times T_{C_2}(x_2),$$

$$N_C(x) = N_{C_1}(x_1) \times N_{C_2}(x_2).$$

Proposition C.3. (Clarke [43], Loewen [97]). Let $C \subseteq \mathbb{R}^n$ be a closed set, and $\nabla d_C(x)$ exist and be different from zero. Then, if x lies outside C, the set C exactly contains a unique closest point c at which the minimum distance to x is attained, such that

$$\nabla d_C(x) = \frac{x - c}{|x - c|}.$$

Theorem C.2. (Clarke [43]). Let $f = g \circ F$, where $F : \mathbb{R}^n \to \mathbb{R}^m$ is Lipschitz near x, and $g : \mathbb{R}^m \to \mathbb{R}$ is Lipschitz near $F(x)$. Then, f is Lipschitz near x and satisfies:

$$\partial f(x) \subset \operatorname{co} \{\partial g(F(x)) \partial F(x)\}.$$

If g is strictly differentiable at $F(x)$, then equality holds and co is needless.

Appendix D

Measure Theory and Integration

D.1 Algebra and $\sigma-$algebra of Sets

Definition D.1. *(Algebra). Let X be an arbitrary non-empty set, and $\Omega(X)$ be a collection of subsets of X. Then, $\Omega(X)$ is called an algebra if satisfies*

(i) $\emptyset, X \in \Omega(X)$*, where* $\emptyset$ *is the empty set.*

(ii) $A \in \Omega(X) \Longrightarrow A^c \in \Omega(X)$*, where* A^c *is the complement of* A *defined by*

$$A^c = \{a \in X \mid a \notin A\}.$$

(iii) $A, B \in \Omega(X) \Longrightarrow A \cup B \in \Omega(X)$.

Definition D.2. *($\sigma-$algebra). An algebra $\Omega(X)$ is said to be a $\sigma-$algebra (sigma-algebra) if satisfies the additional condition:*
For any sequence $\{A_n\} \subset \Omega(X) \Longrightarrow \bigcup_{n=1}^{\infty} A_n \in \Omega(X)$.

Note that, the intersections of this sequence will also belong to $\Omega(X)$ *i.e.*,

$$\{A_n\} \subset \Omega(X) \Longrightarrow \bigcap_{n=1}^{\infty} A_n \in \Omega(X).$$

Definition D.3. *(Borel $\sigma-$algebra). Let X be a metric space. The Borel $\sigma-$algebra of X is defined to be $\sigma-$algebra generated by all open subsets of X, and is denoted by $\mathcal{B}(X)$. Elements of $\mathcal{B}(X)$ are said to be Borel measurable set.*

D.2 Measures

We say that the measure $\mu(\cdot)$ is finite additive if, for any family $S_1, \ldots, S_n \in \Omega$ of disjoint sets, we have

$$\mu\left(\bigcup_{j=1}^{n} S_j\right) = \sum_{j=1}^{n} \mu(S_j).$$

We say that $\mu(\cdot)$ is finite subadditive if, for any family $S_1, \ldots, S_n \in \Omega$ of disjoint sets, we have

$$\mu\left(\bigcup_{j=1}^{n} S_j\right) \leq \sum_{j=1}^{n} \mu(S_j).$$

The measure $\mu(\cdot)$ is monotone if, for any $S_1, S_2 \in \Omega$, with $S_1 \subset S_2$, we have

$$\mu(S_1) \leq \mu(S_2),$$

and is countable subadditivity if, for any sequence $\{S_n\} \subset \Omega$, we have

$$\mu\left(\bigcup_{n\in\mathbb{N}} S_n\right) \leq \sum_{n\in\mathbb{N}} \mu(S_n).$$

Lemma D.1. *A measure $\mu(\cdot)$ on a $\sigma-$algebra Ω of subset of a set X is finite additive, finite subadditive, monotone, countable subadditive, and*

$$\mu(S_2 \setminus S_1) = \mu(S_2) - \mu(S_1),$$

for all $S_1, S_2 \in \Omega$, $S_1 \subset S_2$, with $\mu(S_1) < \infty$.

The proof this Lemma can be found, for instance, in Yeh [154].

Definition D.4. *Consider the measure $\mu\colon \mathcal{B}(X) \to [0, \infty]$ and let $S \in \mathcal{B}(X)$. The measure $\mu(\cdot)$ is called an outer regular measure if*

$$\mu(S) = \inf\{\mu(U) \mid S \subseteq U \textit{ and } U \textit{ is open}\},$$

and is called an inner regular measure if

$$\mu(S) = \sup\{\mu(K) \mid K \subseteq S \textit{ and } K \textit{ is compact}\}.$$

A measure is called a regular if it is both outer and inner regular. A Radon measure is an inner regular Borel measure.

Definition D.5. *Let (X, Ω) and (Y, Ψ) be two measurable spaces. A map $f\colon X \to Y$ is said to be measurable if, for all $A \in \Psi$, the set $f^{-1}(A) \in \Omega$. If Y is a metric space and $\Psi = \mathcal{B}(Y)$, f is called a Borel functional. If $Y = \mathbb{R}$, we call f a Borel function.*

D.3 Integration

Definition D.6. *(Riemann integral) Let $f\colon [a,b] \to \mathbb{R}$ be a bounded function defined on a closed interval $[a,b]$, consider any partition $\mathcal{P}\colon a = x_0 < x_1 < \cdots < x_n = b$, let $\delta x_i = x_i - x_{i-1}$, and we define the upper and lower sum, respectively, associated with partition $\mathcal{P}$ as follows*

$$\begin{aligned} U_{\mathcal{P}} &= \sum_{i=1}^{n} M_i \delta x_i, \\ L_{\mathcal{P}} &= \sum_{i=1}^{n} m_i \delta x_i, \end{aligned}$$

where

$$\begin{aligned} M_i &= \sup\{f(x)\colon x_{i-1} < x \le x_i\}, \\ m_i &= \inf\{f(x)\colon x_{i-1} < x \le x_i\}. \end{aligned}$$

Then, we define the upper and lower Riemann integral of f over $[a,b]$, respectively, as following

$$\begin{aligned} \overline{\int_a^b} f(x)dx &= \inf_{\mathcal{P}} U_{\mathcal{P}}, \\ \underline{\int_a^b} f(x)dx &= \sup_{\mathcal{P}} L_{\mathcal{P}}. \end{aligned}$$

We say that f is Riemann integrable on $[a,b]$, denoted by $\int_a^b f(x)dx$, if the infimum of upper sums through all partitions of $[a,b]$ is equal to the supremum of all lower sums through all partitions of $[a,b]$, i.e.,

$$\overline{\int_a^b} f(x)dx = \underline{\int_a^b} f(x)dx = \int_a^b f(x)dx.$$

Definition D.7. *(Simple function) A function is simple if its range is a finite set. Let ψ is a simple function represented by*

$$\psi = \sum_{i=1}^{n} a_i \chi_{E_i},$$

where a_i are distinct values of ψ and χ_{E_i} is a measurable function called the indicator function of the set E_i given by

$$\chi_{E_i}(x) = \begin{cases} 1, & \text{if } x \in E_i, \\ 0, & \text{if } x \notin E_i, \end{cases}$$

such that $E_i = \psi^{-1}(a_i)$. *Conversely, any expression of this form, where* a_i *need not be distinct and* E_i *not necessarily* $\psi^{-1}(a_i)$ *also defines a simple function.*

Definition D.8. *Let* (X, μ) *be a measure space,* $\mu(\cdot)$ *is Lebesgue measure. Then, The Lebesgue integral over* $D \subset X$ *of a measurable valued simple function* ψ *is defined by*

$$\int_D \psi d\mu = \int_D \sum_{i=1}^{n} a_i \chi_{E_i} d\mu = \sum_{i=1}^{n} a_i \mu(E_i).$$

The quantity on the right represents the sum of the areas below the graph of $\psi(\cdot)$.

Definition D.9. *Let* $f(\cdot)$ *be a bounded measurable function, defined on a set* D *of finite measure such that the upper and lower Lebesgue integral, respectively, defined by*

$$U = \sup \left\{ \int_D \psi d\mu \mid \psi \text{ is a simple function, and } \psi \leq f \right\},$$

$$L = \inf \left\{ \int_D \psi d\mu \mid \psi \text{ is a simple function, and } \psi \geq f \right\}.$$

If $U = L$*, we call that* f *is Lebesgue integrable, and*

$$L = \int_D f d\mu = U$$

where $\int_D \psi d\mu$ is stated before.

Definition D.10. *Let* (X, Ω, μ) *be a measure space. A measurable function* $f \colon X \to \mathbb{R}$ *is called Lebesgue integrable on* $Y \in \Omega$ *with respect to* μ*, if the non-negative function* $|f| = f^+ + f^-$ *satisfies*

$$\int_Y |f| \, d\mu < \infty,$$

and its integral on Y *is defined as*

$$\int_Y f d\mu = \int_Y f^+ d\mu - \int_Y f^- d\mu$$

where f^+, f^- *are said to be the positive and negative parts of the function* f*, defined by*

$$\begin{aligned} f^+(x) &= \max(f(x), 0), \\ f^-(x) &= \max(-f(x), 0). \end{aligned}$$

For more details and properties on Lebesgue integral (see, *e.g.*, Carter and Van Brunt [38], Thomson [143], and Yeh *et al.* [154]).

D.4 Useful Concepts

A function $f(t)$, defined on the closed interval $[a, b]$ is said to be (i) increasing if

$$f(t_1) \leq f(t_2), \quad \text{for } t_1 < t_2,$$

(ii) strictly increasing if

$$f(t_1) < f(t_2), \quad \text{for } t_1 < t_2,$$

(iii) decreasing if

$$f(t_1) \geq f(t_2), \quad \text{for } t_1 < t_2,$$

(iv) strictly decreasing if

$$f(t_1) > f(t_2), \quad \text{for } t_1 < t_2,$$

(v) monotonic or monotone (strictly monotonic) if increasing and decreasing (strictly increasing and strictly decreasing) functions.

Definition D.11. (Strong local minimum) An admissible process (x^*, u^*) is a strong local minimizer for an optimal-control problem if, for $\varepsilon > 0$, it minimizes the cost over all other admissible processes (x,u) such that

$$|x(t) - x^*(t)| \leq \varepsilon, \quad \forall t \in [a, b].$$

Definition D.12. (Weak convergence) Let X be a normed linear vector space and X^* be the dual space of X. A sequence $\{x_n\}$ is called converge weakly to $x \in X$ if for all $x^* \in X^*$, we have $\langle x_n, x^* \rangle \to \langle x, x^* \rangle$, we can write, $x_n \to x$ weakly.

Definition D.13. *(Weak* convergence) Let X be a normed linear vector space and X^* be the dual space of X. A sequence $\{x_n^*\}$ in X^* is called converge Weak* (weak-star) to $x^* \in X^*$ if $\langle x, x_n^* \rangle \to \langle x, x^* \rangle$ for all $x \in X$, we can write, $x_n^* \to x^*$ Weak*.*

Bibliography

[1] J. Abello, P. M. Pardalos, and M. G. Resende, *Handbook of massive data sets*, Vol. 4, Springer, 2013.

[2] O. P. Agrawal, *Formulation of Euler–Lagrange equations for fractional variational problems*, Journal of Mathematical Analysis and Applications **272** (2002), no. 1, 368–379.

[3] ______, *A general formulation and solution scheme for fractional optimal control problems*, Nonlinear Dynamics **38** (2004), no. 1-4, 323–337.

[4] ______, *Fractional variational calculus in terms of Riesz fractional derivatives*, Journal of Physics A: Mathematical and Theoretical **40** (2007), no. 24, 6287.

[5] O. P. Agrawal and D. Baleanu, *A Hamiltonian formulation and a direct numerical scheme for fractional optimal control problems*, Journal of Vibration and Control **13** (2007), no. 9-10, 1269–1281.

[6] O. P. Agrawal, O. Defterli, and D. Baleanu, *Fractional optimal control problems with several state and control variables*, Journal of Vibration and Control **16** (2010), no. 13, 1967–1976.

[7] B. Ahmad and S. K. Ntouyas, *Fractional differential inclusions with fractional separated boundary conditions*, Fractional Calculus and Applied Analysis **15** (2012), no. 3, 362–382.

[8] H. M. Ali, F. L. Pereira, and S. M. A. Gama, *A new approach to the Pontryagin maximum principle for nonlinear fractional optimal control problems*, Mathematical Methods in the Applied Sciences (2016).

[9] R. Almeida and D. F. Torres, *Fractional variational calculus for nondifferentiable functions*, Computers & Mathematics with Applications **61** (2011), no. 10, 3097–3104.

[10] P. Arena, R. Caponetto, L. Fortuna, and D. Porto, *Nonlinear noninteger order circuits and systems: an introduction*, Vol. 38, World Scientific, 2000.

[11] A. Arutyunov, *Optimality conditions: Abnormal and degenerate problems*, Vol. 526, Springer, 2000.

[12] A. Arutyunov, V. Dykhta, and F. L. Pereira, *Necessary conditions for impulsive nonlinear optimal control problems without a priori normality assumptions*, Journal of Optimization Theory and applications **124** (2005), no. 1, 55–77.

[13] A. V. Arutyunov, D. Y. Karamzin, and F. L. Pereira, *Maximum principle in problems with mixed constraints under weak assumptions of regularity*, Optimization **59** (2010), no. 7, 1067–1083.

[14] ______, *The maximum principle for optimal control problems with state constraints by R.V. Gamkrelidze: revisited*, Journal of Optimization Theory and Applications **149** (2011), no. 3, 474–493.

[15] A. V. Arutyunov, D. .Y Karamzin, and F. L. Pereira, *On a generalization of the impulsive control concept: controlling system jumps*, Discrete Contin. Dyn. Syst **29** (2011), no. 2, 403–415.

[16] A. V. Arutyunov, D. Y. Karamzin, and F. L. Pereira, *R.V. Gamkrelidze's maximum principle for optimal control problems with bounded phase coordinates and its relation to other optimality conditions*, Doklady Mathematics **83** (2011), no. 1, 131–135.

[17] ______, *State constraints in impulsive control problems: Gamkrelidze-like conditions of optimality*, Journal of Optimization Theory and Applications **166** (2015), no. 2, 440–459.

[18] ______, *Conditions for the absence of jumps of the solution to the adjoint system of the maximum principle for optimal control problems with state constraints*, Proceedings of the Steklov Institute of Mathematics **292** (2016), no. 1, 27–35.

[19] A. V. Arutyunov, D. Y. Karamzin, F. L. Pereira, and G. N. Silva, *Investigation of regularity conditions in optimal control problems with geometric mixed constraints*, Optimization **65** (2016), no. 1, 185–206.

[20] A. V. Arutyunov and F. L. Pereira, *Second-order necessary conditions for optimal impulsive control*, IFAC Proceedings Volumes (IFACP–apersOnline), 2005, pp. 941–941.

[21] ______, *Second-order necessary optimality conditions for problems without a priori normality assumptions*, Mathematics of Operations Research **31** (2006), no. 1, 1–12.

[22] A.V. Arutyunov, D. Y. Karamzin, and F. L. Pereira, *Pontryagin's maximum principle for constrained impulsive control problems*, Nonlinear Analysis: Theory, Methods & Applications **75** (2012), no. 3, 1045–1057.

[23] T. M. Atanackovic and B. Stankovic, *On a numerical scheme for solving differential equations of fractional order*, Mechanics Research Communications **35** (2008), no. 7, 429–438.

[24] M. Axtell and M. E. Bise, *Fractional calculus application in control systems*, Proceedings of the IEEE 1990 National Aerospace and Electronics Conference (NAECON), (1990), pp. 563–566.

[25] R. L. Bagley and J. Torvik, *Fractional calculus-A different approach to the analysis of viscoelastically damped structures*, AIAA Journal **21** (1983), no. 5, 741–748.

[26] R. L. Bagley and P. J. Torvik, *A theoretical basis for the application of fractional calculus to viscoelasticity*, Journal of Rheology **27** (1983), no. 3, 201–210.

[27] D. Baleanu, O. Defterli, and O. P. Agrawal, *A central difference numerical scheme for fractional optimal control problems*, Journal of Vibration and Control **15** (2009), no. 4, 583–597.

[28] D. Baleanu, S. Rezapour, and H. Mohammadi, *Some existence results on nonlinear fractional differential equations*, Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences **371** (2013), no. 1990, 20120144.

[29] H. Behncke, *Optimal control of deterministic epidemics*, Optimal Control Applications and Methods **21** (2000), no. 6, 269–285.

[30] M. Benchohra, S. Djebali, and S. Hamani, *Boundary-value problems for differential inclusions with Riemann–Liouville fractional derivative*, Nonlinear Oscillations **14** (2011), no. 1, 6–20.

[31] M. Benchohra, J. Henderson, S. K. Ntouyas, and A. Ouahab, *Existence results for fractional functional differential inclusions with infinite delay and applications to control theory*, Fractional Calculus and Applied Analysis **11** (2008), no. 1, 35–56.

[32] E. Bishop and R. R. Phelps, *The support functionals of a convex set*, Proceedings of Symposia in Pure Mathematics, 1963, pp. 27–35.

[33] L. Bourdin and D. Idczak, *A fractional fundamental lemma and a fractional integration by parts formula–applications to critical points of Bolza functionals and to linear boundary value problems*, Advances in Differential Equations **20** (2015), no. 3/4, 213–232.

[34] H. E. Bray, *Elementary properties of the Stieltjes integral*, Annals of Mathematics **20** (1919), no. 3, 177–186.

[35] A. E. Bryson, *Applied optimal control: optimization, estimation and control*, CRC Press, 1975.

[36] A. J. Calderón, B. M. Vinagre, and V. Feliu, *Fractional order control strategies for power electronic buck converters*, Signal Processing **86** (2006), no. 10, 2803–2819.

[37] R. Caponetto, G. Dongola, L. Fortuna, and I. Petráš, *Modeling and control applications*, World Scientific, 2010.

[38] M. Carter and B. Van Brunt, *The Lebesgue-Stieltjes integral*, Springer, 2000.

[39] A. Cernea, *On a fractional differential inclusion with boundary conditions*, Studia Universitatis Babes-Bolyai, Mathematica **3** (2010), 105–113.

[40] P. Chen and S. M. Islam, *Optimal control models in finance*, Springer, 2005.

[41] C. W. Clark, *Mathematical Bioeconomics: The optimal Management of Sustainable Resources*, Wiley-Interscience, New York, 1990.

[42] F. Clarke, *The maximum principle in optimal control, then and now*, Control and Cybernetics **34** (2005), 709–722.

[43] F. H. Clarke, *Optimization and nonsmooth analysis*, Wiley, New York, 1983.

[44] F. H. Clarke, Y. S. Ledyaev, R. J. Stern, and P. R. Wolenski, *Nonsmooth analysis and control theory*, Springer, 1998.

[45] Y. Cohen, *Applications of control theory in Ecology: Proceedings of the Symposium on Optimal Control theory held at the State University of New York, Syracuse, New York, August 10–16, 1986*, Vol. 73, Springer Science & Business Media, 2013.

[46] C. Corduneanu, *Principles of differential and integral equations*, Vol. 295, American Mathematical Society, 2008.

[47] M. Dalir and M. Bashour, *Applications of fractional calculus*, Applied Mathematical Sciences **4** (2010), no. 21, 1021–1032.

[48] S. Das, *Functional fractional calculus*, Springer Science & Business Media, 2011.

[49] B. E. Davis and D. J. Elzinga, *The solution of an optimal control problem in financial modeling*, Operations Research **19** (1971), no. 6, 1419–1433.

[50] G. De Barra, *Measure theory and integration*, Elsevier, 2003.

[51] E. C. de Oliveira and J. A. Tenreiro, *A review of definitions for fractional derivatives and integrals*, Mathematical Problems in Engineering **1** (2014), no. 2, 238459.

[52] T. C. Doehring, A. D. Freed, E. O. Carew, and I. Vesely, *Fractional order viscoelasticity of the aortic valve cusp: an alternative to quasilinear viscoelasticity*, Journal of Biomechanical Engineering **127** (2005), no. 4, 700–708.

[53] A. Y. Dubovitskii and A. A. Milyutin, *Extremum problems with constraints*, Sov. Math., Dokl. **4** (1963), 452–455.

[54] M. O. Efe, *Fractional order systems in industrial automation - a survey*, IEEE Transactions on Industrial Informatics **7** (2011), no. 4, 582–591.

[55] I. Ekeland, *On the variational principle*, Journal of Mathematical Analysis and Applications **47** (1974), no. 2, 324–353.

[56] A. M. A. El-Sayed and A. G. Ibrahim, *Set-valued integral equations of fractional-orders*, Applied Mathematics and Computation **118** (2001), no. 1, 113–121.

[57] V. Feliu-Batlle, R. R. Perez, and L. S. Rodriguez, *Fractional robust control of main irrigation canals with variable dynamic parameters*, Control Engineering Practice **15** (2007), no. 6, 673–686.

[58] G. B. Folland, *Real analysis: modern techniques and their applications*, John Wiley & Sons, 2013.

[59] S. L. Fraga and F. L. Pereira, *Hamilton-Jacobi-Bellman equation and feedback synthesis for impulsive control*, Automatic Control, IEEE Transactions on **57** (2012), no. 1, 244–249.

[60] R. V. Gamkrelidze, *Optimal processes with bounded phase coordinates*, Izv. Akad. Nauk SSSR. Ser. Mat. **24** (1960), 315–356.

[61] A. Golbabai and K. Sayevand, *Analytical treatment of differential equations with fractional coordinate derivatives*, Computers & Mathematics with Applications **62** (2011), no. 3, 1003–1012.

[62] R. Gorenflo and F. Mainardi, *Essentials of fractional calculus*, Mathematical Physics and Stochastics Center (2000), 226–227.

[63] J. Gregory and C. Lin, *Constrained optimization in the calculus of variations and optimal control theory*, Springer, 1992.

[64] E. V. Grigorieva, E. N. Khailov, and E. I. Kharitonova, *Optimal control of pollution stock through ecological interaction of the manufacturer and the state*, Revista de Matematica: Teoria y Aplicaciones **18** (2011), 77–109.

[65] R. E. Gutiérrez, J. M. Rosário, and J. Tenreiro Machado, *Fractional order calculus: basic concepts and engineering applications*, Mathematical Problems in Engineering **2010** (2010), 375858.

[66] S. E. Hamamci, *An algorithm for stabilization of fractional-order time delay systems using fractional-order PID controllers*, IEEE Transactions on Automatic Control **52** (2007), no. 10, 1964–1969.

[67] B. A. Hawkins and H. V. Cornell, *Theoretical approaches to biological control*, Cambridge University Press, 1999.

[68] M. R. Hestenes, *Calculus of variations and optimal control theory*, Wiley, New York, 1966.

[69] R. Hilfer, *Threefold introduction to fractional derivatives*, Anomalous Transport: Foundations and Applications (2008), 17–73.

[70] R. Hilfer, P. L. Butzer, U. Westphal, J. Douglas, W. R. Schneider, G. Zaslavsky, T. Nonnemacher, A. Blumen, and B. West, *Applications of fractional calculus in physics*, Vol. 5, World Scientific, 2000.

[71] K. Hoyt, B. Castaneda, M. Zhang, P. Nigwekar, P. A. di Sant'Agnese, J. V. Joseph, J. Strang, D. J. Rubens, and K. J. Parker, *Tissue elasticity properties as biomarkers for prostate cancer*, Cancer Biomarkers **4** (2008), no. 4, 5, 213–225.

[72] M. Ichise, Y. Nagayanagi, and T. Kojima, *An analog simulation of non-integer order transfer functions for analysis of electrode processes*, Journal of Electroanalytical Chemistry and Interfacial Electrochemistry **33** (1971), no. 2, 253–265.

[73] Z. D. Jelicic and N. Petrovacki, *Optimality conditions and a solution scheme for fractional optimal control problems*, Structural and Multidisciplinary Optimization **38** (2009), no. 6, 571–581.

[74] Y. Jin, Y. Q. Chen, and D. Xue, *Time-constant robust analysis of a fractional order [proportional derivative] controller*, Control Theory & Applications, IET **5** (2011), no. 1, 164–172.

[75] G. Jumarie, *On the representation of fractional Brownian motion as an integral with respect to* $(dt)^\alpha$, Applied Mathematics Letters **18** (2005), no. 7, 739–748.

[76] ———, *Modified Riemann-Liouville derivative and fractional Taylor series of nondifferentiable functions further results*, Computers & Mathematics with Applications **51** (2006), no. 9, 1367–1376.

[77] ———, *Table of some basic fractional calculus formulae derived from a modified Riemann-Liouville derivative for non-differentiable functions*, Applied Mathematics Letters **22** (2009), no. 3, 378–385.

[78] ———, *Fractional differential calculus for non-differentiable functions: Mechanics, geometry, stochastics, information theory*, Lap Lambert Academic, 2013.

[79] E. Jung, S. Lenhart, and Z. Feng, *Optimal control of treatments in a two-strain tuberculosis model*, Discrete and Continuous Dynamical Systems Series B **2** (2002), no. 4, 473–482.

[80] R. Kamocki, *On the existence of optimal solutions to fractional optimal control problems*, Applied Mathematics and Computation **235** (2014), 94–104.

[81] ———, *Pontryagin maximum principle for fractional ordinary optimal control problems*, Mathematical Methods in the Applied Sciences **37** (2014), no. 11, 1668–1686.

[82] ———, *Variational methods for a fractional Dirichlet problem involving Jumarie derivative*, Mathematical Problems in Engineering **2015** (2015), 248517.

[83] R. Kamocki and C. Obczyński, *On fractional differential inclusions with the Jumarie derivative*, Journal of Mathematical Physics **55** (2014), no. 2, 022902.

[84] D. Y. Karamzin, V. A. Oliveira, F. L. Pereira, and G. N. Silva, *On some extension of optimal control theory*, European Journal of Control **20** (2014), no. 6, 284–291.

[85] ———, *Parameter uncertainty in state constrained optimal control problems*, PHYSCON 2015, Istanbul, Turkey, 19–22 August, 2015.

[86] N. Karapetyants and A. Ginsburg, *Lebesgue points for fractional integral*, Integral Transforms and Special Functions **3** (1995), no. 3, 201–210.

[87] A. A. A. Kilbas, H. M. Srivastava, and J. J. Trujillo, *Theory and applications of fractional differential equations*, Vol. 204, Elsevier Science Limited, 2006.

[88] Y. Kobayashi, H. Watanabe, T. Hoshi, K. Kawamura, and M. G. Fujie, *Viscoelastic and nonlinear liver modeling for needle insertion simulation*, Soft Tissue Biomechanical Modeling for Computer Assisted Surgery **11** (2012), 41–67.

[89] C. G. Koh and J. M. Kelly, *Application of fractional derivatives to seismic analysis of base-isolated models*, Earthquake Engineering & Structural Dynamics **19** (1990), no. 2, 229–241.

[90] P. Lanusse, R. Malti, and P. Melchior, *Crone control system design toolbox for the control engineering community: tutorial and case study*, Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences **371** (2013), no. 1990, 20120149.

[91] P. Lanusse and J. Sabatier, *PLC implementation of a crone controller*, Fractional Calculus and Applied Analysis **14** (2011), no. 4, 505–522.

[92] N. Laskin and G. Zaslavsky, *Nonlinear fractional dynamics on a lattice with long range interactions*, Physica A: Statistical Mechanics and its Applications **368** (2006), no. 1, 38–54.

[93] M. J. Lazo and D. F. Torres, *The Dubois–Reymond fundamental lemma of the fractional calculus of variations and an Euler–Lagrange equation involving only derivatives of Caputo*, Journal of Optimization Theory and Applications **156** (2013), no. 1, 56–67.

[94] S. Lenhart and J. T. Workman, *Optimal control applied to biological models*, CRC Press, 2007.

[95] S. Y. Lin, *Generalized Gronwall inequalities and their applications to fractional differential equations*, Journal of Inequalities and Applications **2013** (2013), no. 1, 549.

[96] Z. Liu and B. Zeng, *Existence and controllability for fractional evolution inclusions of Clarke subdifferential type*, Applied Mathematics and Computation **257** (2015), 178–189.

[97] P. D. Loewen, *Optimal control via nonsmooth analysis*, American Mathematical Society, 1993.

[98] A. Lotfi, M. Dehghan, and S. A. Yousefi, *A numerical technique for solving fractional optimal control problems*, Computers & Mathematics with Applications **62** (2011), no. 3, 1055–1067.

[99] A. Luo and V. Afraimovich, *Long-range interactions, stochasticity and fractional dynamics: Dedicated to George M. Zaslavsky (1935–2008)*, Springer Science & Business Media, 2011.

[100] Y. Luo and Y. Chen, *Stabilizing and robust fractional order PI controller synthesis for first order plus time delay systems*, Automatica **48** (2012), no. 9, 2159–2167.

[101] Y. Luo, T. Zhang, B. Lee, C. Kang, and Y. Chen, *Fractional–order proportional derivative controller synthesis and implementation for hard-disk-drive servo system*, IEEE Transactions on Control Systems Technology **22** (2014), no. 1, 281–289.

[102] E. Macé, I. Cohen, G. Montaldo, R. Miles, M. Fink, and M. Tanter, *In vivo mapping of brain elasticity in small animals using shear wave imaging*, IEEE Transactions on Medical Imaging **30** (2011), no. 3, 550–558.

[103] J. T. Machado, V. Kiryakova, and F. Mainardi, *Recent history of fractional calculus*, Communications in Nonlinear Science and Numerical Simulation **16** (2011), no. 3, 1140–1153.

[104] F. Mainardi, *Fractional relaxation-oscillation and fractional diffusion-wave phenomena*, Chaos, Solitons & Fractals **7** (1996), no. 9, 1461–1477.

[105] N. Makris, G. F. Dargush, and M. C. Constantinou, *Dynamic analysis of viscoelastic-fluid dampers*, Journal of Engineering Mechanics **121** (1995), no. 10, 1114–1121.

[106] A. B. Malinowska, T. Odzijewicz, and D. F. Torres, *Advanced methods in the fractional calculus of variations*, Springer, 2015.

[107] A. B. Malinowska and D. F. Torres, *Introduction to the fractional calculus of variations*, World Scientific, 2012.

[108] G. M. Mittag-Leffler, *Sur la nouvelle fonction $E_\alpha(x)$*, C.R. Acad. Sci. Paris **137** (1903), 554–558.

[109] C. A. Monje, Y. Chen, B. M. Vinagre, D. Xue, and V. Feliu-Batlle, *Fractional–order systems and controls: fundamentals and applications*, Springer Science & Business Media, 2010.

[110] B. S. Mordukhovich, *Maximum principle in the problem of time optimal response with nonsmooth constraints*, Journal of Applied Mathematics and Mechanics **40** (1976), no. 6, 960–969.

[111] ———, *Generalized differential calculus for nonsmooth and set-valued mappings*, Journal of Mathematical Analysis and Applications **183** (1994), no. 1, 250–288.

[112] D. Mozyrska and D. F. M. Torres, *Modified optimal energy and initial memory of fractional continuous–time linear systems*, Signal Processing **91** (2011), no. 3, 379–385.

[113] J. D. Munkhammar, *Riemann-Liouville fractional derivatives and the Taylor-Riemann series*, UUDM Project Report **7** (2004), 1–18.

[114] I. P. Natanson, *Theory of functions of a real variable*, Frederick Ungar, New York, 1957.

[115] L. W. Neustadt, *An abstract variational theory with applications to a broad class of optimization problems. II. Applications*, SIAM Journal on Control **5** (1967), no. 1, 90–137.

[116] Z. M. Odibat and N. T. Shawagfeh, *Generalized Taylor's formula*, Applied Mathematics and Computation **186** (2007), no. 1, 286–293.

[117] S. Okada and W. J. Ricker, *Criteria for determining compactness of the integration operator associated with a vector measure*, Tatra Mt. Math. Publ **24** (2002), no. 175, 175–188.

[118] A. Oustaloup, O. Cois, P. Lanusse, P. Melchior, X. Moreau, and J. Sabatier, *The CRONE aproach: Theoretical developments and major applications*, Fractional Differentiation and its Applications **2** (2006), no. 1, 324–354.

[119] F. L. Pereira and G. N. Silva, *Necessary conditions of optimality for vector-valued impulsive control problems*, Systems & Control Letters **40** (2000), no. 3, 205–215.

[120] ———, *Stability for impulsive control systems*, Dynamical Systems **17** (2002), no. 4, 421–434.

[121] ———, *A maximum principle for infinite time asymptotically stable impulsive dynamic control systems*, IFAC Proceedings Volumes (IFAC–PapersOnline) (2010), 1326–1331.

[122] ———, *Necessary conditions of optimality for state constrained infinite horizon differential inclusions*, 50th IEEE Conference on Decision and Control and European Control Conference (CDC–ECC), (2011), pp. 6717–6722.

[123] F. L. Pereira, G. N. Silva, and V. Oliveira, *Invariance for impulsive control systems*, Automation and Remote Control **69** (2008), no. 5, 788–800.

[124] I. Petras, *Fractional–order nonlinear systems: modeling, analysis and simulation*, Springer Science & Business Media, 2011.

[125] I. Podlubny, *Fractional differential equations: an introduction to fractional derivatives, fractional differential equations, to methods of their solution and some of their applications*, Vol. 198, Academic Press, 1998.

[126] ———, *Fractional–order systems and $PI^{\lambda}D^{\mu}$-controllers*, IEEE Transactions on Automatic Control **44** (1999), no. 1, 208–214.

[127] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko, *The mathematical theory of optimal processes*, Wiley Interscience, New York, 1962.

[128] S. Pooseh, R. Almeida, and D. F. M. Torres, *Fractional order optimal control problems with free terminal time*, Journal of Industrial and Management Optimization **10** (2014), no. 2, 363–381.

[129] J. C. Prajapati, *Certain properties of Mittag–Leffler function with argument x^{α}, $\alpha > 0$*, Italian Journal of Pure and Applied Mathematics **30** (2013), 411–416.

[130] M. Renardy, W. J. Hrusa, and J. A. Nohel, *Mathematical problems in viscoelasticity*, A Merican Mathematical Society, New York, 1987.

[131] H. S. Rodrigues, M. T. T. Monteiro, and D. F. M. Torres, *Optimal control and numerical software: an overview*, preprint arXiv:1401.7279 (2014).

[132] B. Ross, S. G. Samko, and E. R. Love, *Functions that have no first order derivative might have fractional derivatives of all orders less than one*, Real Analysis Exchange **20** (1994), no. 2, 140–157.

[133] Y. A. Rossikhin and M. V. Shitikova, *Applications of fractional calculus to dynamic problems of linear and nonlinear hereditary mechanics of solids*, Applied Mechanics Reviews **50** (1997), no. 1, 15–67.

[134] H. A. Salem and M. Cichoń, *On solutions of fractional order boundary value problems with integral boundary conditions in banach spaces*, Journal of Function Spaces and Applications **2013** (2013), 428094.

[135] S. G. Samko, A. A. Kilbas, and O. I. Marichev, *Fractional integrals and derivatives, theory and applications*, Gordon and Breach, Amsterdam 1993, Engl. Transl. from the Russian Edition, 1987.

[136] S. P. Sethi and G. L. Thompson, *Optimal control theory: applications to management science and economics*, Springer, 2000.

[137] G. W. Swan, *Applications of optimal control theory in biomedicine*, CRC Press, 1984.

[138] V. E. Tarasov, *Universal electromagnetic waves in dielectric*, Journal of Physics: Condensed Matter **20** (2008), no. 17, 175223.

[139] ———, *Fractional dynamics: applications of fractional calculus to dynamics of particles, fields and media*, Springer Science & Business Media, 2011.

[140] M. Tavazoei, *Time response analysis of fractional-order control systems: A survey on recent results*, Fractional Calculus and Applied Analysis **17** (2014), no. 2, 440–461.

[141] M. S. Tavazoei, M. Haeri, S. Jafari, S. Bolouki, and M. Siami, *Some applications of fractional calculus in suppression of chaotic oscillations*, IEEE Transactions on Industrial Informatics **55** (2008), no. 11, 4094–4101.

[142] M. E. Taylor, *Measure theory and integration*, American Mathematical Society, 2006.

[143] B. S. Thomson, *Theory of the integral*, ClassicalRealAnalyis. com, 2012.

[144] B. S. Thomson, J. B. Bruckner, and A. M. Bruckner, *Elementary real analysis*, ClassicalRealAnalysis. com, 2008.

[145] A. Torokhti and P. Howlett, *Dynamic programming and its application to optimal control*, Vol. 81, Elsevier, 1971.

[146] C. Tricaud and Y. Chen, *An approximate method for numerically solving fractional order optimal control problems of general form*, Computers & Mathematics with Applications **59** (2010), no. 5, 1644–1655.

[147] C. C. Tseng and S. L. Lee, *Digital IIR integrator design using recursive Romberg integration rule and fractional sample delay*, Signal Processing **88** (2008), no. 9, 2222–2233.

[148] D. Valério and J. Sá da Costa, *Non-integer order control of a flexible robot*, Proceedings of the IFAC workshop on fractional differentiation and its applications, Bordeaux, France, 2004, pp. 520–525.

[149] B. M. Vinagre, Y. Q. Chen, and I. Petráš, *Two direct Tustin discretization methods for fractional-order differentiator/integrator*, Journal of the Franklin Institute **340** (2003), no. 5, 349–362.

[150] R. B. Vinter, *Optimal control*, Springer, 2000.

[151] R. B. Vinter and F. L. Pereira, *A maximum principle for optimal processes with discontinuous trajectories*, SIAM journal on control and optimization **26** (1988), no. 1, 205–229.

[152] B. West, M. Bologna, and P. Grigolini, *Physics of fractal operators*, Springer Science & Business Media, 2012.

[153] H. Ye, J. Gao, and Y. Ding, *A generalized Gronwall inequality and its application to a fractional differential equation*, Journal of Mathematical Analysis and Applications **328** (2007), no. 2, 1075–1081.

[154] J. Yeh, *Real analysis: theory of measure and integration*, World Scientific, 2006.

[155] S. A. Yousefi, A. Lotfi, and M. Dehghan, *The use of a Legendre multiwavelet collocation method for solving the fractional optimal control problems*, Journal of Vibration and Control **17** (2011), no. 13, 2059–2065.

[156] M. I. Zelikin and V. F. Borisov, *Theory of chattering control: with applications to astronautics, robotics, economics, and engineering*, Springer Science & Business Media, 2012.

[157] B. Zheng, *Some new Gronwall-Bellman-type inequalities based on the modified Riemann-Liouville fractional derivative*, Journal of Applied Mathematics **2013** (2013), 341706.